



# Zeros of the Selberg zeta function for symmetric infinite area hyperbolic surfaces

Mark Pollicott<sup>1</sup> · Polina Vytnova<sup>1</sup>

Received: 21 May 2018 / Accepted: 18 August 2018  
© The Author(s) 2018

## Abstract

In the present paper we give a simple mathematical foundation for describing the zeros of the Selberg zeta functions  $Z_X$  for certain very symmetric infinite area surfaces  $X$ . For definiteness, we consider the case of three funneled surfaces. We show that the zeta function is a complex almost periodic function which can be approximated by complex trigonometric polynomials on large domains (in Theorem 4.2). As our main application, we provide an explanation of the striking empirical results of Borthwick (Exp Math 23(1):25–45, 2014) (in Theorem 1.5) in terms of convergence of the affinely scaled zero sets to standard curves  $\mathcal{C}$ .

**Keywords** Selberg zeta function · Infinite area surface · Transfer operator

**Mathematics Subject Classification** 11M36 · 37C30 · 42A75

## 1 Introduction

The Selberg zeta function  $Z_X$  associated to a compact Riemann surface  $X$  with negative Euler characteristic and without boundary is a well known and much studied complex function. It is a function of a single complex variable defined in terms of the lengths  $\ell(\gamma)$  of the primitive closed geodesics  $\gamma$  on the surface by analogy with the Riemann zeta function in number theory.

**Definition 1.1** We can formally define the Selberg zeta function by

$$Z_X(s) = \prod_{n=0}^{\infty} \prod_{\substack{\gamma = \text{prime} \\ \text{closed geodesics}}} \left(1 - e^{-(s+n)\ell(\gamma)}\right), \quad s \in \mathbb{C}, \quad (1.1)$$

where the product is taken over all primitive closed geodesics  $\gamma$  on  $X$ .

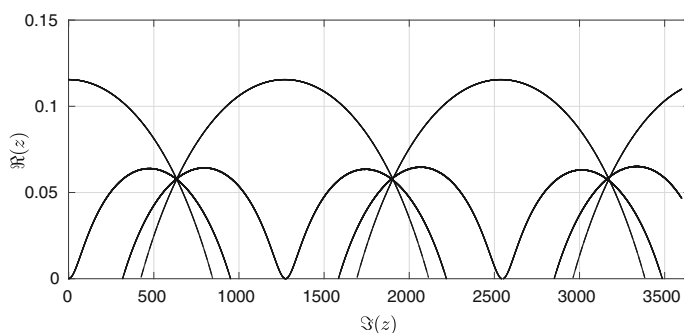
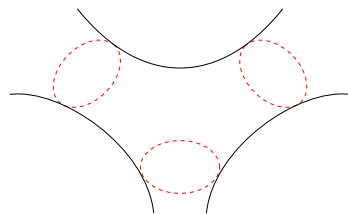
---

✉ Mark Pollicott  
masdbl@warwick.ac.uk

Polina Vytnova  
P.Vytnova@warwick.ac.uk

<sup>1</sup> Warwick University, Coventry, UK

**Fig. 1** An artist's impression of a symmetric three funnelled surface. The three dashed closed geodesics are assumed to be of equal length  $2b$



**Fig. 2** The zeros of the associated zeta function  $Z_X$  in a critical strip. The individual zeros are so close in the plot that it creates the illusion that they lie on well defined smooth curves

It was shown by A. Selberg in 1956 that for such surfaces the zeta function  $Z_X$  has an analytic extension to the entire complex plane and that the non-trivial zeros can be described in terms of the spectrum of the Laplace–Beltrami operator [29] and the Selberg trace formula.

**Theorem 1.2** (Selberg) *Let  $X$  be a compact Riemann surface with negative Euler characteristic and without boundary. Then the function  $Z_X$  has a simple zero at  $s = 1$  and for any zero  $s$  in the critical strip  $0 < \Re(s) < 1$  we have that either  $s \in [0, 1]$  is real, or  $\Re(s) = \frac{1}{2}$ .*

A good reference for this standard result is Chapter 2 of the book of Hejhal [14].

In the case of infinite area surfaces the situation is somewhat different, since the original trace formulae approach of Selberg no longer applies. However, it still follows from the dynamical method of Ruelle that providing the surface  $X$  appears as a quotient space of the upper half-plane  $\mathbb{H}^2$  by a convex cocompact Fuchsian group  $\Gamma$  then the analogous zeta function  $Z_X$  still has an analytic extension to the entire complex plane [26, 28]. Unfortunately, this approach provides little effective information on the location of the zeros, other than that the first zero at  $s = \delta \in (0, 1)$ , the Hausdorff dimension of the limit set (of the Fuchsian group).

In pioneering experimental work, D. Borthwick has studied the location of the zeros for the zeta function in specific examples of infinite area surfaces [6]. The plot in Fig. 2 is fairly typical for the distribution of zeros in the critical strip for a symmetric three funnelled surface  $X_b$ , for large  $b$ , where each of the three simple closed geodesics corresponding to a funnel has the same length  $2b$ .<sup>1</sup>

The analysis of the Selberg zeta function is via the action of the Fuchsian group on the boundary of hyperbolic space and the basic method dates back to work of the first author over

<sup>1</sup> This is an infinite area surface  $X_b$  whose compact core  $K_b$  corresponds to a sphere with three disks removed (called a “pair of pants”) and geodesic boundary components corresponding to the three unique simple closed geodesics of length  $2b$  around each funnel (as represented in Fig. 1).

twenty five years ago [27]. However, it is only with the advent of superior computational resources, and the ingenuity of those that employ them, that the striking features seen in Fig. 2, for example, have been revealed. A contemporary personal computer allows one to study symmetric three funneled surface with the length of the three defining closed geodesics at least 8 without much difficulty, and this turns out to be a sufficiently general case. We refer the reader to [33] for more references to the existing literature.

The main object of study in the present work is the zero set of the function  $Z_X$ :

$$\mathcal{S}_X := \{s \in \mathbb{C} \mid Z_X(s) = 0\}. \quad (1.2)$$

Carefully studying plots of a few thousands zeros of zeta functions associated to symmetric 3-funnelled surfaces in a domain  $|s| \leq 3000$  and  $\Re(s) > 0$  for example, one observes that there are certain similarities in a way the zeros are arranged. We formulate our observations in terms of the theory of almost periodic functions.

**Definition 1.3** We say that  $\tau \in i\mathbb{R}$  is a  $\varepsilon$ -translation of a discrete set  $\{A\} \subset \mathbb{C}$  if there is a bijection  $\{A\} \xrightarrow{\varphi} \{A + \tau\}$  such that  $d(a, \varphi(a)) \leq \varepsilon$  for all  $a \in \{A\}$ . We say that a discrete set is almost periodic in the sense of Krein–Levin, if for any  $\varepsilon > 0$  there exists a relatively dense<sup>2</sup> set of  $\varepsilon$ -translations.

**Informal qualitative observations** Let  $X_b$  be the three funnelled surface defined by three simple closed geodesics of equal length<sup>3</sup>  $2b$ . Then for a sufficiently large  $b$ .

- O1 : The set of zeros  $\mathcal{S}_{X_b}$  appears to be an almost periodic set in the sense of Krein–Levin, with relatively dense set of translations  $\tau = \{i(\pi k e^b + \varepsilon_k) \mid k \in \mathbb{N}\}$ , where  $\varepsilon_k = o(e^{-b/2})$  as  $b \rightarrow +\infty$ .
- O2 : The set of zeros  $\mathcal{S}_{X_b}$  appears to lie on a few distinct curves, which seem to have a common intersection point at  $\frac{\delta}{2} + i\frac{\pi}{2}e^b$ , as  $b \rightarrow +\infty$ .

It is well known that the zeta-function can be well approximated by a sequence of finite exponential sums [17,28], and therefore the first property is to be expected: the zero set of a finite exponential sum is almost periodic (cf. [19], Appendix 1). The main difficulty here is to identify the set of translations.

The second property is more mysterious, and this will follow from a specific approximation of the zeta function by an exponential sum of 12 terms on a large, but bounded, domain, which grows with  $b$  exponentially quickly.

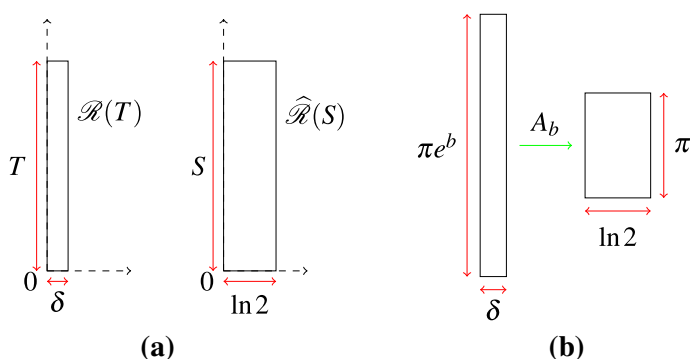
In order to provide a rigorous proof, we need to estimate the error term of approximation of the zeta function by exponential sums as  $b \rightarrow +\infty$ . This is done in Theorem 4.2. Subsequently using properties of the zeta function we simplify the exponential sum, which gives the best approximation, and obtain a function whose zero set belongs to four sinusoids.

According to an old result by McMullen [21] the largest real zero  $\delta$  asymptotically behaves like  $\frac{\ln 2}{b}$  as  $b \rightarrow +\infty$ , and it defines the width of the critical strip. Therefore, in the limit  $b \rightarrow +\infty$  the zero set converges to imaginary axis. However, a suitable affine rescaling allows one to see the pattern of zeros for large values of  $b$ . A natural choice for rescaling factors is the approximate period of the pattern (in the imaginary direction) and approximate reciprocal of the width of the critical strip (in the real direction) (Fig. 3).

**Notation 1.4** We will be using the following.

<sup>2</sup> The set  $\mathcal{T} \in i\mathbb{R}$  is relatively dense if there exists  $l > 0$  such that for any interval  $I \subset i\mathbb{R}$  of the length  $l$  we have that  $I \cap \mathcal{T} \neq \emptyset$ .

<sup>3</sup> This normalization makes formulae in subsequent calculations shorter.



**Fig. 3** **a** The strips  $\mathcal{R}(T)$  and  $\widehat{\mathcal{R}}(S)$ ; **b** By renormalizing the strip  $\mathcal{R}(\pi e^b)$  to  $\widehat{\mathcal{R}}(\pi)$  we can compare the zeros of zeta functions for different  $b$ , as  $b$  tends to infinity

(a) A compact part of the critical strip of height  $T > 0$  which we denote by

$$\mathcal{R}(T) = \{s \in \mathbb{C} \mid 0 \leq \Re(s) \leq \delta \text{ and } |\Im(s)| \leq T\};$$

and a compact part of the normalized critical strip of height  $S > 0$  which we denote by

$$\widehat{\mathcal{R}}(S) = \{s \in \mathbb{C} \mid 0 \leq \Re(s) \leq \ln 2 \text{ and } |\Im(s)| \leq S\}.$$

(b) We denote the rescaled set of zeros by

$$\widehat{\mathcal{S}}_{X_b} := \left\{ \sigma b + i e^{-b} t \mid \sigma + it \in \mathcal{S}_X \right\}.$$

where evidently,  $0 < \Re(\widehat{\mathcal{S}}_{X_b}) \leq \ln 2$ . We now introduce a family of four sinusoidal curves approximating  $\widehat{\mathcal{S}}_{X_b}$  as  $b \rightarrow +\infty$ .

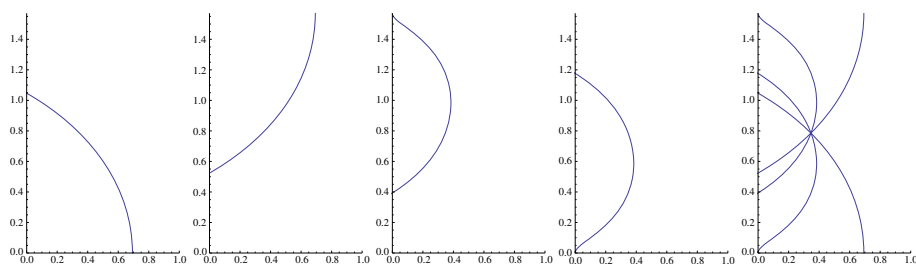
(c) Let  $\mathcal{C} = \bigcup_{j=1}^4 \mathcal{C}_j$ , where

$$\begin{aligned} \mathcal{C}_1 &= \left\{ \frac{1}{2} \ln |2 - 2 \cos(t)| + it \mid t \in \mathbb{R} \right\}; \\ \mathcal{C}_2 &= \left\{ \frac{1}{2} \ln |2 + 2 \cos(t)| + it \mid t \in \mathbb{R} \right\}; \\ \mathcal{C}_3 &= \left\{ \frac{1}{2} \ln \left| 1 - \frac{1}{2} e^{2it} - \frac{1}{2} e^{it} \sqrt{4 - 3e^{2it}} \right| + it \mid t \in \mathbb{R} \right\}; \\ \mathcal{C}_4 &= \left\{ \frac{1}{2} \ln \left| 1 - \frac{1}{2} e^{2it} + \frac{1}{2} e^{it} \sqrt{4 - 3e^{2it}} \right| + it \mid t \in \mathbb{R} \right\}. \end{aligned}$$

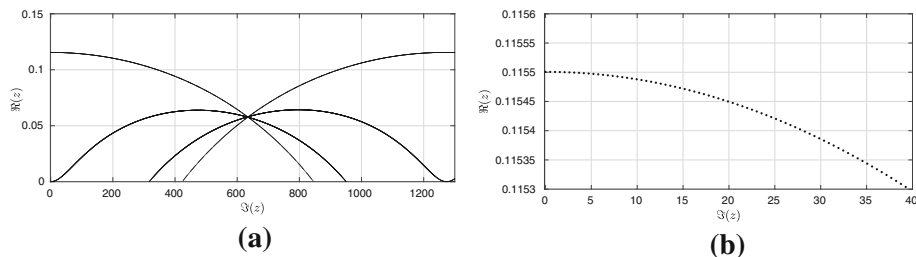
Note that the curve  $\mathcal{C}$  in Fig. 4 looks similar to empirical plots in Fig. 5. The apparent intersections in Fig. 2 correspond to the intersections of  $\mathcal{C}_j$ :

$$\bigcap_{j=1}^4 \mathcal{C}_j = \left\{ \frac{\ln 2}{2} + i\pi \left( \frac{1}{2} + k \right) \right\}, \quad k \in \mathbb{Z}.$$

We can now formally state the approximation result, which provides an explanation for the Observations.



**Fig. 4** Plots of the curves  $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4$  and their union  $\mathcal{C}$



**Fig. 5** **a** The zeros of the  $Z_{14}$  (see §4 for definition) approximating  $Z_{X_6}$ ; **b** A zoomed version in a neighbourhood of  $\delta$  showing individual zeros. The distance between imaginary parts of consecutive zeros is approximately  $\frac{\pi}{6}$

**Theorem 1.5** *The sets  $\widehat{\mathcal{F}}_{X_b}$  and  $\mathcal{C}$  are close in the Hausdorff metric  $\text{dist}_H$  on a large part of the strip  $0 < \Re(s) < \log 2$ . More precisely, there exists  $\varkappa > 0$  such that*

$$\text{dist}_H(\widehat{\mathcal{F}}_{X_b} \cap \widehat{\mathcal{R}}(e^{\varkappa b}), \mathcal{C} \cap \widehat{\mathcal{R}}(e^{\varkappa b})) = O\left(\frac{1}{\sqrt{b}}\right), \text{ as } b \rightarrow +\infty.$$

The theorem implies that every rescaled zero  $s \in \widehat{\mathcal{F}}_{X_b} \cap \widehat{\mathcal{R}}(e^{\varkappa b})$  belongs to a neighbourhood of  $\mathcal{C}$  which is shrinking as  $b \rightarrow \infty$ . On the other hand, the rescaled zeros are so close, that the union of their shrinking neighbourhoods contains  $\mathcal{C}$ .

**Remark 1.6** Although an explicit estimate on the distance between zeros and curves  $O\left(\frac{1}{\sqrt{b}}\right)$  is a bonus, the most significant feature of this result is that the height  $e^{\varkappa b}$  of the rescaled strip  $\widehat{\mathcal{R}}(e^{\varkappa b})$  is larger than the period of the curves  $\mathcal{C}$ , and it corresponds to a part of the original strip of height  $e^{(1+\varkappa)b}$  which allows us to find a set of  $\varepsilon$ -translations.

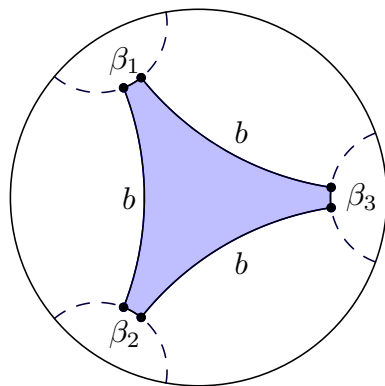
**Remark 1.7** A similar analysis can be carried out for a punctured torus and for less symmetric surfaces. However, in most cases the modul<sup>4</sup> of the exponential sum, approximating the zeta function with suitable accuracy, has more generators, the curves containing zero set can be defined only implicitly, and the set of  $\varepsilon$ -translations doesn't have such a simple form.

There is an interesting conjecture due to Jakobson and Naud which we include here for context (see [15], Conjecture 1.1, p. 354).

**Conjecture 1.8** (Jakobson–Naud). *There are only finitely many zeros in the half plane  $\Re(s) > \frac{\delta}{2}$  and this is the largest half-plane with this property (i.e., for any  $\varepsilon > 0$  there are infinitely many zeros in the half-plane  $\Re(s) > \frac{\delta}{2} - \varepsilon$ ).*

<sup>4</sup> The minimal subalgebra of  $\mathbb{R}$  containing all multipliers.

**Fig. 6** Three geodesics of reflection in Poincaré disk with pairwise distance  $b$ . The compact connected subset  $K_b$  of  $X_b$  bounded by the three simple closed geodesics is a double cover for the hyperbolic hexagon illustrated



We also note that there are interesting empirical investigations on the (Ruelle) zeta function in the case of Sinai billiards [31].

This project has had a long gestation period, having begun after the first author heard the original empirical results of D. Borthwick presented at a conference on Quantum Chaos in Roscoff in June of 2013. We are grateful to him for sharing his original Matlab code with us. We are grateful to F. Bykov for writing a new program. A preliminary announcement of these results was made by the first author at the conference “Spectral problems for hyperbolic dynamical systems” held in Bordeaux in May of 2014. A preliminary version of this paper was written during authors’ visit to ICERM, Brown University. We would like to thank ICERM for the hospitality.

## 2 The zeta function and closed geodesics

We begin this section with the some basic background on the surface  $X_b$ . Any surface with constant curvature  $\kappa = -1$  has universal covering space the unit disc  $\mathbb{D}^2 = \{z = x + iy : |z| < 1\}$  equipped with the usual Poincaré metric  $ds^2 = 4 \frac{dx^2 + dy^2}{(1 - x^2 - y^2)^2}$ . In particular,  $(\mathbb{D}^2, ds)$  is a simply connected Riemann surface with constant negative curvature  $-1$  and  $X_b$  can be viewed as a double oriented cover of a quotient  $\mathbb{D}^2 / \Gamma$  by a discrete subgroup of isometries  $\Gamma$  (i.e., a Fuchsian group).

There are many different choices of generators for  $\Gamma$ . Because of the natural symmetries of  $X_b$ , it is convenient to choose a presentation of the associated Fuchsian group in terms of three reflections (as in [21], for example). More precisely, we can fix a value  $0 < \theta \leq \frac{\pi}{3}$  and consider the Fuchsian group  $\Gamma = \Gamma_\theta := \langle R_1, R_2, R_3 \rangle$  generated by reflections  $R_1, R_2, R_3$  in three disjoint equidistant geodesics  $\beta_1, \beta_2, \beta_3$ , with end points  $e^{(\frac{2\pi}{3}j \pm \theta)i} \in \partial \mathbb{D}^2$ ,  $j = 1, 2, 3$ , respectively (cf. Fig. 6).

Although the three individual generators are orientation reversing the resulting quotient surface  $X_b = \mathbb{D}^2 / \Gamma$  is an oriented infinite area surface. Both the infinite area surface and the corresponding compact surface  $K_b$  with boundary consisting of three geodesics of the length  $2b$ , share the same closed geodesics, and thus have the same zeta function. Furthermore, the compact surface  $K_b$  (sometimes known as a pair of pants) is precisely the double cover for

the hyperbolic hexagon in Fig. 6, (see Exposé 3, [10]).<sup>5</sup> The values  $b$  and  $\alpha$  can be related using a simple hyperbolic geometry calculation.

Moreover, since  $X_b$  is a negatively curved surface it is a classical result of E. Cartan that we can associate a unique closed geodesic  $\gamma$  to each conjugacy class in the fundamental group  $\pi_1(X_b)$  (see Theorem 2.2 of Chapter of [8]; or Chapter 6 of [4]).

**Notation 2.1** We shall denote by  $\ell(\gamma)$  length of  $\gamma$  in the hyperbolic metric of the surface. We denote by  $\omega(\gamma) \in 2\mathbb{N}$  the word length, that is the (even) number of generators  $R_1, R_2, R_3$  required to represent the conjugacy class corresponding to  $\gamma$  in  $\Gamma$ . Geometrically, the word length corresponds to the period of the associated cutting sequence, i.e. the sequence of reflections corresponding to the sides of the hexagons crossed consecutively by the geodesic [30] and whose period corresponds to the total number of edges crossed.

**Remark 2.2** Therefore, since  $K_b$  is a double cover for a hyperbolic hexagon (cf. Fig. 6) every closed geodesic must traverse each of the two copies of the hexagon consecutively the same number of times and thus the word length is necessarily even.

## 2.1 Analytic approximation

Our starting point for understanding the properties of the zeta function is the following important result of D. Ruelle from 1976, stating that the infinite product (1.1) defines an analytic function.

**Theorem 2.3** (after Ruelle) Let  $\delta > 0$  be the largest real zero for  $Z_{X_b}$ . In the notation introduced above, the infinite product (1.1) converges to a non-zero analytic function for  $\Re(s) > \delta$  and extends as an analytic function to  $\mathbb{C}$ .

**Proof** The convergence follows from more general results on Axiom A flows which we can apply to the geodesic flow restricted to the recurrent part [25]. The analyticity follows from applying ideas from the work of Ruelle [28], see also [25] for more details.  $\square$

The Selberg zeta function, initially defined by (1.1) in terms of lengths of infinitely many closed geodesics, is not an object which can be easily computed numerically. A computer can only deal with a finite set of geodesics. Naturally, a fundamental question arises: how to choose the geodesics to obtain a good approximation to the infinite product?

The approach of Ruelle provides an approximation of the zeta function  $Z_{X_b}$  by finite exponential sums, whose multipliers and coefficients depend on all geodesics corresponding to the word length less than  $n$ , which make it suitable for numerical experiments.

We begin by considering a more general function in two complex variables

$$Z_{X_b}(s, z) = \prod_{n=0}^{\infty} \prod_{\substack{\gamma=\text{primitive} \\ \text{closed geodesic}}} \left(1 - z^{\omega(\gamma)} e^{-(s+n)\ell(\gamma)}\right), \quad (2.1)$$

which converges for  $|z|$  sufficiently small and  $\Re(s)$  sufficiently large. Taking  $z = 1$  we recover the original zeta function  $Z_{X_b}(s) = Z_{X_b}(s, 1)$ . We follow Ruelle [28] in re-writing the infinite product as a series

<sup>5</sup> This is because the reflections  $R_1, R_2, R_3$  reverse orientation. Composing pairs of reflections  $R_1 R_2$  and  $R_1 R_3$ , say, would give orientation preserving boundary identifications on the fundamental domain of  $X_b$  and generate a free group isomorphic to  $\pi_1(X_b)$  as in [6], but at the expense of the natural symmetry.

$$Z_{X_b}(s, z) = 1 + \sum_{n=1}^{\infty} a_n(s) z^n, \quad (2.2)$$

by taking the Taylor expansion in  $z$  about 0. It is then easy to see that  $a_n(s)$  is defined in terms of finitely many closed geodesics with word lengths at most  $n$ . In fact this series converges to a bianalytic function for both  $z, s \in \mathbb{C}$  as shown by the next Theorem (which is stronger than Theorem 2.3).

**Theorem 2.4** (Ruelle [28]) *Using the notation introduced above, there exists  $C = C(s) > 0$  and  $0 < \alpha < 1$  such that  $|a_n| \leq C\alpha^{n^2}$  and thus the series (2.2) converges. In particular, we can deduce that  $Z_{X_b}(s, z)$  is analytic in both variables.*

Theorem 2.3 is a corollary of the stronger Theorem 2.4. We next make an easy observation.

**Lemma 2.5** *The odd coefficients vanish i.e.,  $a_1 = a_3 = \dots = 0$ .*

**Proof** It was observed in Remark 2.2 that  $\omega(\gamma) \in 2\mathbb{N}$  in the infinite product (2.1) and thus only the even terms  $a_2, a_4, \dots$  can be non-zero in (2.2).  $\square$

We can formally rewrite the zeta function (2.1) as

$$Z_{X_b}(s, z) = \exp \left( - \sum_{m=1}^{\infty} \frac{z^m}{m} \sum_{\omega(\gamma)=m} \frac{e^{-s\ell(\gamma)}}{1 - e^{-\ell(\gamma)}} \right) = \exp \left( - \sum_{m=1}^{\infty} \frac{b_m(s)}{m} z^m \right) \quad (2.3)$$

where

$$b_m(s) = \begin{cases} \sum_{\omega(\gamma)=m} \frac{e^{-s\ell(\gamma)}}{1 - e^{-\ell(\gamma)}} & \text{if } m \text{ is even;} \\ 0 & \text{if } m \text{ is odd;} \end{cases} \quad (2.4)$$

and then expand the exponential as a power series and obtain coefficients  $a_n$  comparing (2.2) with (2.3), since these series converge provided  $\Re(s) > 1$ . In particular, we can easily check that the first three non-zero terms are:

$$\begin{aligned} a_2(s) &= -\frac{1}{2} b_2(s); \\ a_4(s) &= -\frac{1}{4} \left( b_4(s) - \frac{b_2^2(s)}{2} \right); \\ a_6(s) &= -\frac{1}{6} \left( b_6(s) + \frac{1}{4} b_2(s) b_4(s) - \frac{b_2^3(s)}{8} \right); \end{aligned}$$

and in general,

$$a_n(s) = -\frac{1}{n} \sum_{j=0}^{n-2} a_j(s) b_{n-j}(s).$$

Combining the latter with (2.4), we deduce that each of the coefficients  $a_n(s)$  is an exponential sum whose multipliers are sums of the lengths of several closed geodesics of the total word length  $n$ .



**Remark 2.6** It is important to know the number of closed geodesics of word length  $2n$  for practical applications. We can count them by counting periodic cutting sequences. It's easy to see that the number of cutting sequences of period  $2n$ , associated to closed geodesics, i.e. satisfying the additional condition  $\xi_1 \neq \xi_{2n}$ , satisfies the recurrence relation  $p_n = 3 \cdot 4^{n-1} + p_{n-1}$ ,  $p_1 = 6$ . Hence there are exactly  $p_n = 4^n + 2$  closed geodesics of the word length  $\omega(\gamma) = 2n$ .

Straightforward approximations to  $Z_{X_b}$  obtained by evaluating the series (2.2) at  $z = 1$  allow one to compute zeros one-by-one numerically using the Newton method. In order to find explicit curves, we need a further simplification. Another approximation to  $Z_{X_b}$  can be obtained by replacing the lengths of closed geodesics  $\ell(\gamma)$  in (1.1) by suitable close approximations. This will result in a function with zero set on  $\mathcal{C}$ . Afterwards, we shall show that both approximations are sufficiently close to each other.

**Remark 2.7** Applying general results cf. [19] Appendix 1, on zeros of exponential sums we also deduce that the zero set of a finite sum  $1 + \sum_{k=1}^n a_k(s)$  belongs to a strip parallel to the imaginary axis, and that the difference between imaginary parts of consecutive zeros is approximately  $\frac{\pi}{b}$ . This was proved by Weich [33] for small values of  $s \in \mathcal{S}_{X_b}$ , without relying on the theory of almost periodic functions, but instead using “symmetry reduction” first suggested in [9].

## 2.2 Geometric approximation

The following simple trick allows us to reduce the problem of locating zeros of the double infinite product (1.1) to a problem of locating zeros of a single infinite product. Namely, consider the related Ruelle zeta function  $\zeta(s)$  defined by

$$\zeta(s) = \prod_{\substack{\gamma = \text{primitive} \\ \text{closed geodesic}}} (1 - e^{-s\ell(\gamma)})^{-1} = \frac{Z_{X_b}(s+1)}{Z_{X_b}(s)}, \quad (2.5)$$

where  $\gamma$  again denotes a primitive closed geodesic of length  $\ell(\gamma)$ . Since  $Z_{X_b}$  is real analytic and non-zero for  $\Re(s) > 1$ , it has poles corresponding to the zeros  $Z_{X_b}(s)$  in the strip  $0 < \Re(s) < \delta$ . The function  $\zeta$  is the exact form of the zeta function studied by Ruelle [28] and is better suited to geometric approximation.

We will be following the approach in [25], a similar argument can be found in [32].

Let  $\Sigma$  be the set of infinite cutting sequences corresponding to geodesics on the 3-funnelled surface  $X_b$ . We consider a function

$$\tilde{r}_n: \Sigma \rightarrow \mathbb{R} \quad \tilde{r}_n(\xi) = \ell(\gamma_{\xi_{[n/2]\xi_{[n/2]+1}}}), \quad (2.6)$$

where  $\gamma_\xi$  is a closed geodesic corresponding to the cutting sequence  $\xi$  and  $\gamma_{\xi_{[n/2]\xi_{[n/2]+1}}}$  is the shortest geodesic segment between the intersections with the sides corresponding to  $R_{\xi_{[n/2]}}$  and  $R_{\xi_{[n/2]+1}}$  which lies on a longer segment passing consecutively through the sides corresponding to the reflections  $R_{\xi_1}, \dots, R_{\xi_{n+1}}$ . In particular, any geodesic is uniquely defined by its cutting sequence and therefore if the word length  $\omega(\gamma_\xi) \leq 2n$  we have

$$\ell(\gamma_\xi) = \sum_{j=0}^{\omega(\gamma)-1} \tilde{r}_n(\sigma^j \xi); \text{ and} \quad (2.7)$$

in general, we have that

$$\ell(\gamma_\xi) = \lim_{n \rightarrow \infty} \sum_{j=0}^{\omega(\gamma)-1} \tilde{r}_n(\sigma^j \xi), \quad (2.8)$$

where  $\sigma: \Sigma \rightarrow \Sigma$  is a shift given by  $\sigma(\xi_n) = (\xi_{n+1})$ .

We know that periodic sequences in  $\Sigma$  are periodic orbits of a subshift of finite type (cf. [23], pp. 11–12). We denote by  $A$  the transition matrix of the subshift corresponding to the encoding by subsequences of length  $n$ . Let  $\xi^1, \dots, \xi^N$  be all subsequences of the sequences in  $\Sigma$  of the length  $n$ . (In the case of  $X_b$  we have that  $N = 3 \cdot 2^{n-1}$ .) We define an  $N \times N$  matrix

$$A_{i,j}^n = \begin{cases} 1, & \text{if } \xi_{k+1}^i = \xi_k^j, \text{ for } k = 1, \dots, n-1 \\ 0, & \text{otherwise.} \end{cases} \quad (2.9)$$

We use the transition matrix  $A^n$  for the subshift to define a family of matrices  $A(s)$  whose elements depend on the length of geodesics segments corresponding to transitions. Namely, we introduce a matrix-valued complex function  $A(s)$  by

$$A: \mathbb{C} \rightarrow \text{Mat}(N, N) \quad A_{i,j}(s) = \exp(-s\tilde{r}_n(\xi)) \cdot A_{i,j}^n, \quad (2.10)$$

where  $\xi = \xi_1^i \dots \xi_n^j$ . Note that  $A(s)$  depends on  $n$  but we omit this.

The following fact is very useful and relatively well-known cf. [23, 25, 32].

**Lemma 2.8** *Using the notation introduced above the following equality holds*

$$\prod_{\substack{\gamma = \text{primitive} \\ \text{closed geodesic}}} (1 - e^{-s\ell(\gamma)})^2 = \lim_{n \rightarrow \infty} \det(I_N - A^2(s)); \quad (2.11)$$

where  $I_N$  is the  $N \times N$  identity matrix.

**Proof** Observe that on the domain of convergence the right hand side

$$\log(\det(I_N - A^2(s))) = - \sum_{j=1}^{\infty} \frac{1}{j} \text{tr}(A^{2j}(s)) = - \sum_{j=1}^{\infty} \frac{1}{j} \sum_{\sigma^{2j}\xi=\xi} e^{-s(\tilde{r}_n(\xi) + \tilde{r}_n(\sigma\xi) + \dots + \tilde{r}_n(\sigma^{2j-1}\xi))}$$

we may rewrite the latter term as

$$= - \sum_{j=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{jm} \sum_{\substack{\sigma^{2j}\xi=\xi \\ \xi \text{ primitive}}} e^{-sm(\tilde{r}_n(\xi) + \tilde{r}_n(\sigma\xi) + \dots + \tilde{r}_n(\sigma^{2j-1}\xi))},$$

where the inner summation is taken over primitive fixed points for  $\sigma^{2j}$

$$= -2 \sum_{j=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{m} \sum_{\substack{\{\xi, \sigma\xi, \sigma^2\xi, \dots, \sigma^{2j-1}\xi\} \\ \text{primitive}}} e^{-sm(\tilde{r}_n(\xi) + \tilde{r}_n(\sigma\xi) + \dots + \tilde{r}_n(\sigma^{2j-1}\xi))},$$

where the inner summation is taken over primitive closed orbits of period  $2j$

$$\begin{aligned} &= -2 \sum_{j=1}^{\infty} \sum_{\substack{\{\xi, \sigma\xi, \sigma^2\xi, \dots, \sigma^{2j-1}\xi\} \\ \text{primitive}}} \log \left( 1 - e^{-s(\tilde{r}_n(\xi) + \tilde{r}_n(\sigma\xi) + \dots + \tilde{r}_n(\sigma^{2j-1}\xi))} \right), \\ &= -2 \log \prod_{j=1}^{\infty} \prod_{\substack{\{\xi, \sigma\xi, \sigma^2\xi, \dots, \sigma^{2j-1}\xi\} \\ \text{primitive}}} \left( 1 - e^{-s(\tilde{r}_n(\xi) + \tilde{r}_n(\sigma\xi) + \dots + \tilde{r}_n(\sigma^{2j-1}\xi))} \right). \end{aligned}$$

Using (2.8) and swapping the limits, we get the result, since there is a bijection between oriented primitive closed geodesics and primitive periodic cutting sequences.  $\square$

The last lemma establishes a connection between the Ruelle zeta function and the determinant of certain matrices. It follows from definition (2.10) that  $\det(I_N - A^2(s))$  is a finite exponential sum as a function of  $s$ . It turns out that in the particular case of  $X_b$ , the zero set of  $\det(I_N - A^2(s))$  for  $N = 6$  is easy to describe and that for  $b$  large this exponential sum is close to  $1 + \sum_{j=1}^{\infty} a_n(s)$ , where  $a_n(s)$  are the coefficients in (2.2).

## 2.3 Computing a geometric approximation

We apply the method described in the previous paragraph in a simple case  $n = 2$ . In other words, we will be computing the length of segments of closed geodesics taking into account only  $n + 1 = 3$  consecutive elements of the cutting sequences, and computing the total length by summing up the length of the pieces. We will show afterwards that this approximation is good enough.

In order to compute  $\tilde{r}_2$  as defined by (2.6), we observe the following simple fact:

**Lemma 2.9** *Consider a regular hyperbolic hexagon whose even sides are of length  $b > 1$  and whose odd sides are of length  $\varepsilon_b$  then*

$$\varepsilon_b = 2e^{-b/2} + e^{-b} + O(e^{-3b/2}) \quad \text{as } b \rightarrow +\infty.$$

**Proof** We recall that (cf. [3], Theorem 7.19.2)

$$\cosh \varepsilon_b (\sinh b)^2 = \cosh b + (\cosh b)^2.$$

The result follows by expanding the both sides in  $e^{-b}$  as  $b \rightarrow +\infty$  and comparing the expansions.  $\square$

Now we are ready to describe  $\tilde{r}_2: \Sigma \rightarrow \mathbb{R}^+$ .

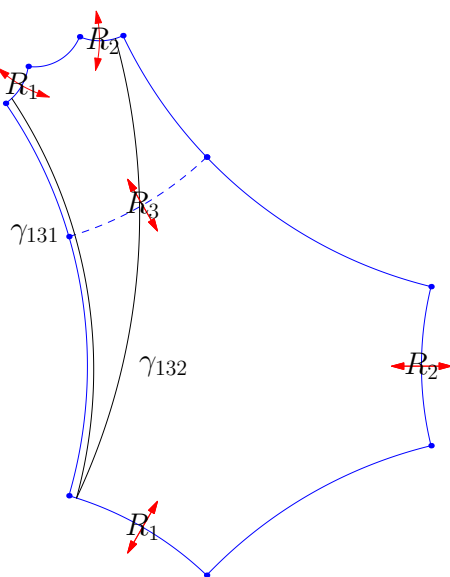
**Lemma 2.10** *Assume that  $\xi \in \Sigma$  is a cutting sequence of period  $2n$  given by*

$$\xi = \xi_1 \xi_2 \xi_3 \cdots \xi_{2n-1} \xi_{2n} \xi_1 \cdots,$$

where  $\xi_k \neq \xi_{k+1}$  for  $1 \leq k \leq 2n$  and  $\xi_{2n} \neq \xi_1$ . Then  $\tilde{r}_2(\xi) = b + c(\xi)e^{-b} + O(e^{-2b})$  as  $b \rightarrow +\infty$ , where

$$c(\xi) = \begin{cases} 0, & \text{if } \xi_1 = \xi_3; \\ 1, & \text{otherwise.} \end{cases}$$

**Fig. 7** A geodesic segment  $\gamma_{131}$ , running along the side of two hexagons when  $R_{\xi_1} = R_{\xi_3}$ ; and a geodesic segment  $\gamma_{132}$ , traversing two hexagons when  $R_{\xi_{n-1}} \neq R_{\xi_{n+1}}$



**Proof** We need to estimate the length of the shortest of the segments of closed geodesics passing through  $\beta_{\xi_1}$ ,  $\beta_{\xi_2}$ , and  $\beta_{\xi_3}$  enclosed between the intersections with  $\beta_{\xi_1}$  and  $\beta_{\xi_2}$ .

Let us first consider the case when  $\xi_1 = \xi_3$ , see geodesic segment  $\gamma_{131}$  in Fig. 7 for example. It is evident that the shortest geodesic whose cutting sequence has a subsequence  $\xi_1 \xi_2 \xi_1$  is the boundary one. Therefore, we conclude that  $\tilde{r}_2(\xi) = b$ .

Now we consider the case  $\xi_1 \neq \xi_3$ , see the geodesic segment  $\gamma_{132}$  in Fig. 7 for example.

We may denote the point of intersection of the geodesic  $\gamma_{\xi}$  with the side  $R_{\xi_k}$  by  $A_k$ . We may assume that  $A_k$  divides the side of reflection in proportions  $x_k : (1 - x_k)$ , for  $k = 1, \dots, 3$ . Then the geodesic segment  $\gamma_{\xi_1 \xi_2}$  is a side of a right hyperbolic trapezoid with one side of the length  $b$ , and two parallel sides of the length  $x_1 \varepsilon_b$  and  $x_2 \varepsilon_b$ . Similarly, the geodesic segment  $\gamma_{\xi_2 \xi_3}$  is a side of a right hyperbolic trapezoid with one side of the length  $b$ , and two parallel sides of the length  $(1 - x_2) \varepsilon_b$  and  $x_3 \varepsilon_b$ .

Using hyperbolic sine and cosine laws, we deduce the formula for length of the fourth side of the trapezoid:

$$\begin{aligned} f_1(x_1, x_2) &:= \ell(\gamma_{\xi_1 \xi_2}) \\ &= \text{Arcosh}(\cosh b \cdot \cosh(x_1 \varepsilon_b) \cdot \cosh(x_2 \varepsilon_b) - \sinh(x_1 \varepsilon_b) \cdot \sinh(x_2 \varepsilon_b)). \end{aligned}$$

Similarly, for another segment:

$$\begin{aligned} f_2(x_2, x_3) &:= \ell(\gamma_{\xi_2 \xi_3}) = \\ &= \text{Arcosh}(\cosh b \cdot \cosh((1 - x_2) \varepsilon_b) \cdot \cosh(x_3 \varepsilon_b) - \sinh((1 - x_2) \varepsilon_b) \cdot \sinh(x_3 \varepsilon_b)). \end{aligned}$$

Consider the function  $f_{\xi}(x_1, x_2, x_3) = f_1(x_1, x_2) + f_2(x_2, x_3)$ . By definition,  $\tilde{r}_2(\xi) = f_1(x'_1, x'_2)$ , where  $x'_1$  and  $x'_2$  are chosen so that

$$f_{\xi}(x'_1, x'_2, x_3) = \inf_{\xi} \inf_{x_1, x_2, x_3} f_{\xi}(x_1, x_2, x_3). \quad (2.12)$$

Analysing this function we obtain that  $x'_1 = 0$  and  $x'_2 = 0.5$ . We see that

$$f_1(0, 0.5) = \text{Arcosh}(\cosh b \cdot \cosh(0.5 \varepsilon_b)).$$

Finally by a straightforward calculation we can write

$$\begin{aligned}\tilde{r}_2(\xi) &= \ln \left( \cosh b \cosh \frac{\varepsilon b}{2} + \sqrt{\cosh^2 b \cosh^2 \frac{\varepsilon b}{2} - 1} \right) \\ &= \ln \left( 2 \cosh b \cosh \frac{\varepsilon b}{2} + \frac{1}{2} e^{-b} + O(e^{-2b}) \right) \\ &= \ln(2 \cosh b) + 2 \ln \left( \cosh \frac{\varepsilon b}{2} \right) + 2 \ln \left( 1 + \frac{1}{2} e^{-2b} + O(e^{-3b}) \right).\end{aligned}\quad (2.13)$$

We have asymptotic expansions

$$\ln(2 \cosh b) = b + \ln(1 + e^{-2b}) = b + e^{-2b} + O(e^{-4b}) \quad (2.14)$$

and

$$\ln \left( \cosh \frac{\varepsilon b}{2} \right) = \ln \left( \frac{e^{e^{-b}/2 + O(e^{-b})} + e^{-e^{-b}/2 + O(e^{-b})}}{2} \right) = \ln \left( 1 + \frac{e^{-b}}{2} + O(e^{-2b}) \right). \quad (2.15)$$

Substituting (2.14) and (2.15) into (2.13), we conclude  $\tilde{r}_2(\xi) = b + e^{-b} + O(e^{-2b})$  as  $b \rightarrow +\infty$ . This completes the case  $\xi_1 \neq \xi_3$  and proves the Lemma.  $\square$

In “Appendix A” we obtain estimates for the length of closed geodesics of the word length 2, 4, 6, and 8; illustrating Lemma 2.10 and formula (2.8).

Now we can use Lemma 2.10 to compute the matrix  $A(s)$  defined by (2.10) in the case  $n = 2$ . We write the transition matrix  $A$  defined by (2.9), corresponding to 6 subsequences of cutting sequences of length two:  $\xi^1 = \{1, 2\}$ ;  $\xi^2 = \{1, 3\}$ ;  $\xi^3 = \{2, 1\}$ ;  $\xi^4 = \{2, 3\}$ ;  $\xi^5 = \{3, 1\}$ ;  $\xi^6 = \{3, 2\}$ :

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Substituting  $\tilde{r}_2(s)$  into (2.10), we get the matrix function:

$$A(s) = \begin{pmatrix} 0 & 0 & e^{-sb} & e^{-s(b+e^{-b})} & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{-sb} & e^{-s(b+e^{-b})} \\ e^{-sb} & e^{-s(b+e^{-b})} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{-s(b+e^{-b})} & e^{-sb} \\ e^{-s(b+e^{-b})} & e^{-sb} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{-s(b+e^{-b})} & e^{-sb} & 0 & 0 \end{pmatrix} + O(e^{-2b}).$$

Introducing a shorthand notation  $z = e^{-s}e^{-b}$  we rewrite the main term of  $A(s)$  as

$$A(s) = e^{-sb} \begin{pmatrix} 0 & 0 & 1 & z & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & z \\ 1 & z & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & z & 1 \\ z & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & z & 1 & 0 & 0 \end{pmatrix}. \quad (2.16)$$

The simplicity of the matrix  $A(s)$  allows us to study the zero set of the analytic function

$$g(s) = \det(I_6 - A^2(s)) = \det(I_6 - A(s)) \cdot \det(I_6 + A(s)).$$

Since  $\exp(-sb)$  doesn't vanish, we see that  $g(s) = 0$  if and only if

$$\det(\exp(2sb)I_6 - \exp(2sb)A^2(s)) = 0$$

and this inequality holds true if and only if  $\exp(2sb)$  is an eigenvalue of  $\exp(2sb) \cdot A^2(s)$ .

The eigenvalues of the matrix

$$B(z) := \begin{pmatrix} 0 & 0 & 1 & z & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & z \\ 1 & z & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & z & 1 \\ z & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & z & 1 & 0 & 0 \end{pmatrix}^2 \quad (2.17)$$

can be computed explicitly:

$$\begin{aligned} \mu_1(z) &= (z - 1)^2, \\ \mu_2(z) &= (z + 1)^2, \\ \mu_3(z) &= 1 - \frac{z^2}{2} + z \frac{\sqrt{4 - 3z^2}}{2}, \\ \mu_4(z) &= 1 - \frac{z^2}{2} - z \frac{\sqrt{4 - 3z^2}}{2}. \end{aligned}$$

We summarise our finding in the following Lemma.

**Lemma 2.11** *The determinant  $\det(I_6 - A^2(s))$  vanishes if and only if  $\exp(2bs) = \mu_k(\exp(-se^{-b}))$  for some  $k = 1, 2, 3, 4$  where  $\mu_k$  as defined above.*

We will show how this Lemma leads to the construction of the four curves containing zeros after computing the errors in approximations of the zeta function by an exponential sum and the exponential sum by the determinant  $\det(I_6 - A^2(s))$ . We would like to finish this section with the following remark on properties of the matrix  $B$ .

**Remark 2.12** Let us consider the matrix  $B(z)$  defined by (2.17). Then the coefficients of the polynomial

$$\frac{1}{2n} \operatorname{tr}(B^n(z)) = d_{2n}z^{2n} + d_{2n-1}z^{2n-1} + \cdots + d_1z + d_0$$

are given by

$$d_k = \# \left\{ \gamma_\xi : \omega(\gamma_\xi) = 2n \text{ and } \sum_{j=0}^{2n-1} c(\sigma^j \xi) = k \right\}.$$

In particular,  $\frac{1}{2n} \operatorname{tr} B^{2n}(1)$  is equal to the number of closed geodesics of the word length  $2n$ .

### 3 Nuclear operators and analytic functions

Error estimates for approximations of the function  $Z_{X_b}$  are based on the original approach in [28] (and the interpretation in [27]). We begin by recalling some abstract results, essentially due to Grothendieck, on nuclear operators. We then complete the section by relating the length of the boundary geodesics to the contraction on the boundary corresponding to reflections, generating the group  $\Gamma_\theta$ .

#### 3.1 Nuclear operators

The convergence of the series (2.2) in Theorem 2.4 will follow from estimates of Ruelle [28], after Grothendieck [11]. We summarize below the general theory.

Let  $\mathcal{B}$  be a Banach space.

**Definition 3.1** We say that a linear operator  $T : \mathcal{B} \rightarrow \mathcal{B}$  is nuclear if there exist for each  $n \geq 1$

1.  $w_n \in \mathcal{B}$ , with  $\|w_n\|_{\mathcal{B}} = 1$ ;
2.  $v_n \in \mathcal{B}^*$ , with  $\|v_n\|_{\mathcal{B}^*} = 1$ ;
3.  $\lambda_n \in \mathbb{R}$ , with  $0 < \lambda < 1$ ,  $C > 0$  satisfying  $|\lambda_n| \leq C\lambda^n$

such that

$$Tf = \sum_{n=1}^{\infty} \lambda_n w_n v_n(f). \quad (3.1)$$

**Lemma 3.2** (after Grothendieck) *A nuclear operator on a Banach space is trace class, and we can write*

$$\det(I - zT) = \exp\left(-\sum_{n=1}^{\infty} \frac{z^n}{n} \operatorname{tr} T^n\right).$$

where the infinite series on the right hand side converges in a small disk  $|z| < \varepsilon$  and the equality between analytic continuations then holds on  $\mathbb{C}$ .

As the left hand side is an analytic function we may expand it in a power series at  $z = 0$ :

$$\det(I - zT) = 1 + \sum_{n=1}^{\infty} a_n z^n,$$

where

$$a_n = \sum_{j_1 < \dots < j_n} \det([v_{j_k}(w_{j_l})]_{k,l=1}^n) \lambda_{j_1} \dots \lambda_{j_n}. \quad (3.2)$$

Applying estimates of Grothendieck and Ruelle, we obtain an explicit bound.

$$|a_n| \leq C^n n^{n/2} \lambda^{n(n+1)/2}, \quad (3.3)$$

where  $n^{n/2}$  bounds the supremum norm of the matrix<sup>6</sup>  $[v_{j_k}(w_{j_l})]_{k,l=1}^n$  and  $C \in \mathbb{R}$  is a constant.

<sup>6</sup> We follow Ruelle in including the term  $n^{n/2}$  although this can be improved upon by looking at Hilbert spaces of analytic functions. For instance, O. Bandtlow and O. Jenkinson [2] have shown that we can suppress  $n^{n/2}$  by working with Hardy spaces, but then  $\lambda$  would be different, too.

### 3.2 Constructing the Banach space

In computations it will prove more useful to use the equivalent representation of hyperbolic space by the upper half-plane  $\mathbb{H}^2 = \{x + iy : y > 0\}$  with the metric  $ds^2 = \frac{dx^2 + dy^2}{y^2}$ .

Given four points  $z_1 < w_1 < w_2 < z_2$  on the boundary  $\partial\mathbb{H}^2$ , we define the cross ratio by

$$[z_1, w_1, w_2, z_2] = \frac{(z_1 - w_2)(w_1 - z_2)}{(z_1 - w_1)(w_2 - z_2)}.$$

We recall the following classical formula (cf. [3] §7.23).

**Lemma 3.3** *Let  $L_1, L_2$  be two disjoint geodesics in  $\mathbb{H}^2$  with end points  $z_1, z_2$  and  $w_1, w_2$ . The distance  $d(L_1, L_2)$  between  $L_1$  and  $L_2$  satisfies  $[z_1, w_1, w_2, z_2] = \tanh^2(d(L_1, L_2)/2)$ .*

By assumption, the group  $\Gamma = \langle R_1, R_2, R_3 \rangle$  is generated by reflections with respect to three disjoint geodesics, which we denote by  $\beta_1, \beta_2$ , and  $\beta_3$ , respectively. Without loss of generality, we may assume that the geodesic  $\beta_j$  has end points  $e^{\left(\frac{2\pi j}{3} \pm \theta\right)i}$ , for  $j = 1, 2, 3$  and a small real number  $\theta$ . More precisely, by straightforward calculation using Lemma 3.3 we get

**Lemma 3.4** *Let  $\beta_1$  and  $\beta_2$  be two disjoint geodesics in  $\mathbb{D}^2$  with end points  $e^{\left(\pm \frac{2\pi}{3} \pm \theta\right)i}$ . Then  $\sin \theta = \frac{1}{2 \cosh b}$ .*

**Proof** We can apply Lemma 3.3 with  $z_1 = e^{\left(\frac{2\pi}{3} + \theta\right)i}$ ,  $z_2 = e^{\left(\frac{2\pi}{3} - \theta\right)i}$ ,  $w_1 = e^{\left(-\frac{2\pi}{3} + \theta\right)i}$ ,  $w_2 = e^{\left(-\frac{2\pi}{3} - \theta\right)i}$  and  $b = d(L_1, L_2)$ .  $\square$

**Remark 3.5** In the notation and under the hypothesis of the last lemma, we have an asymptotic relation

$$\theta = \frac{1}{2}e^{-b} \left(1 + e^{-2b} + o(e^{-3b})\right) \text{ as } b \rightarrow \infty.$$

To define the Banach space, we fix a small  $\varphi < \theta$  and introduce three additional geodesics  $v_j$  with end points  $e^{\left(\frac{2\pi j}{3} \pm \varphi\right)i}$ ,  $j = 1, 2, 3$ . We may consider the disk  $\mathbb{D}^2$  as a subset of  $\mathbb{C}$  and formally extend the geodesics  $v_j$  to circles  $\bar{v}_j \subset \mathbb{C}$ . Furthermore, let  $\{U_j\}_{j=1}^3$  be three disks in  $\mathbb{C}$  such that  $\partial U_j = \bar{v}_j$  cf. Fig. 8.

The Banach space of bounded analytic functions  $f : \sqcup_{j=1}^3 U_j \rightarrow \mathbb{C}$  on the union  $\sqcup_{i=1}^3 U_i$  we denote by  $\mathcal{B}$ . We supply it with the supremum norm  $\|f\|_{\mathcal{B}} := \|f\|_{\infty}$ .

### 3.3 Transfer operators

We can now define transfer operators  $\mathcal{L}_s$ , acting on the Banach space  $\mathcal{B}$  of bounded analytic functions on  $\sqcup_{j=1}^3 U_j$ .

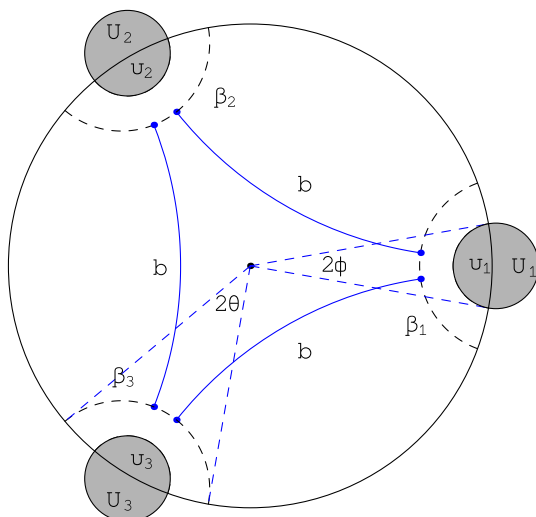
**Definition 3.6** For each  $s \in \mathbb{C}$  we can define

$$(\mathcal{L}_s f)(z) = \sum_{k=1}^3 \chi_{U_k}(z) \sum_{j \neq k} (R'_j(z))^s f(R_j(z)) \quad \text{for } f \in \mathcal{B}, \quad (3.4)$$

where  $\chi_{U_k}$  is the indicator function of  $U_k$ .



**Fig. 8** Three geodesics  $\beta_j$  in  $\mathbb{D}^2$  giving rise to the reflections  $R_j$ ; and three additional geodesics  $v_j$  which are used to define domain of the analytic functions in  $\mathcal{B}$ . The disks  $U_j$  are the shaded regions



We can apply the general theory of nuclear operators to the transfer operators by virtue of the following (compare with [17]).

**Lemma 3.7** *The operator  $\mathcal{L}_s : \mathcal{B} \rightarrow \mathcal{B}$  is nuclear.*

**Proof** We observe that the operators  $f \mapsto f \circ R_j$  are nuclear and  $R'_j(z)^s$  are analytic [28]. Thus  $\mathcal{L}_s$  is nuclear, too.  $\square$

Applying Lemma 3.2 to  $T_s : \mathcal{B} \times \{\pm 1\} \rightarrow \mathcal{B} \times \{\pm 1\}$  defined by  $T_s(w, \epsilon) = (\mathcal{L}_s w, -\epsilon)$  we recover Theorem 2.4, see [28] for details. In particular, by equation 2.4 we can write

$$Z_{X_b}(s, z) = \exp \left( - \sum_{m=1}^{\infty} \frac{z^{2m}}{2m} \text{tr} \mathcal{L}_s^{2m} \right) = \det(I - T_s)$$

where

$$\text{tr} \mathcal{L}_s^{2m} = \sum_{k_1, \dots, k_{2m}} \frac{|(R_{k_1} \cdots R_{k_{2m}})'(x_{k_1, \dots, k_{2m}})|^{-s}}{1 - |(R_{k_1} \cdots R_{k_{2m}})'(x_{k_1, \dots, k_{2m}})|^{-1}}$$

where  $x_{k_1, \dots, k_{2m}}$  is the expanding fixed point for  $R_{k_1} \cdots R_{k_{2m}}$ . This is completely analogous to the approach to transfer operators associated to modular surface in [20].

## 4 Estimating approximation errors

**Notation 4.1** *We denote a partial sum of the series (2.2) by  $Z_n$  ( $n \geq 1$ ):*

$$Z_n(s) := 1 + \sum_{k=1}^n a_k(s). \quad (4.1)$$

*Note that it is an exponential sum of  $\lceil \frac{n}{2} \rceil$  terms, since odd terms vanish  $a_{2k+1} = 0$ .*

Our main approximation result is the following.

**Theorem 4.2** Let  $X_b$  be a symmetric 3-funnelled surface with defining geodesics of length  $2b$ . Then the finite partial sums  $Z_n$  give approximations to  $Z_{X_b}$  on the domain  $\mathcal{R}(T)$  and the remainder is bounded as follows:  $\sup_{\mathcal{R}(T)} |Z_{X_b} - Z_n| \leq \eta(b, n, T)$  where  $T = T(b) = e^{\varkappa b}$  for some constant  $\varkappa > 1$  independent of  $b$  and  $n$ , such that

1. for any  $n \geq 14$  we have  $\eta(b, n, T(b)) = O\left(\frac{1}{\sqrt{b}}\right)$  as  $b \rightarrow \infty$
2. for any  $b \geq 20$  we have  $\eta(b, n, T(b)) = O(e^{-bk_1 n^2})$  as  $n \rightarrow \infty$

for some  $k_1 > 0$  which is independent of  $b$  and  $n$ .

For a fixed  $b$  this theorem estimates the number of terms  $a_n$  needed to uniformly approximate  $Z_{X_b}$  to any given error; at the same time for a given  $n$  this theorem estimates the difference between  $Z_{X_b}$  and  $Z_n$  as  $b \rightarrow \infty$  on an exponentially growing domain.

**Remark 4.3** The constants  $\varkappa$  and  $k_1$  in Theorem 4.2 should satisfy the inequality  $0 < k_1 < \frac{2-\varkappa}{60}$ , although this bound is not sharp. A sharp bound can be obtained using the same argument, but the formulae will be more complicated.

To illustrate Theorem 4.2, we can fix a surface by choosing the length of boundary geodesics  $2b$  and plot the zeros for the approximating trigonometric polynomials  $Z_{2n} = 1 + a_2 + \dots + a_{2n}$  for  $n = 1, 2, \dots, 6$ . For instance, in Fig. 9 zeros of polynomials approximating  $Z_{X_b}$  with  $b = 5$  are shown. The apparent “gaps” in the zeros are due to instability of the Newton method.

**Remark 4.4** In practice, Theorem 4.2 shows that numerical results obtained for  $Z_n$  hold in a domain  $\mathcal{R}(e^{k_1 b})$  for the Selberg zeta function, too. Since in practice  $n$  is bounded above by computational considerations, we may assume that it is fixed. Even with a modern computer, one will not be able to consider  $n > 16$  in a reasonable time.<sup>7</sup> Moreover, in practical applications  $b$  cannot be chosen too large either due to computer restrictions because of accumulation of errors while dealing with small numbers. The coefficients  $b_n$  defined by (2.4) involve a sum of  $4^n + 2$  terms of the order  $\exp(-2snb)$  with  $0 < \Re(s) < 0.25$ , say. The bound  $4^n + 2$  is equal to number of closed geodesics of the word length  $2n$ , see Remark 2.6.

## 4.1 Proof of the approximation result

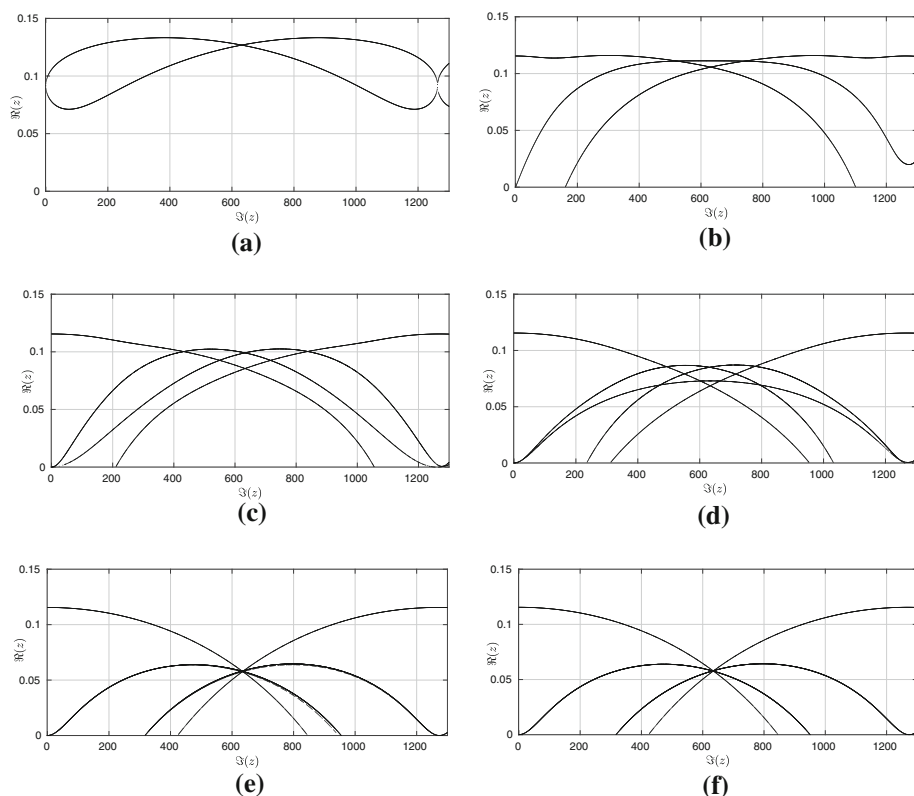
In this section we give a proof of Theorem 4.2. We will need the following simple technical estimate.

**Lemma 4.5** Let  $x_n$  be a sequence of real numbers satisfying  $|x_n| \leq \exp(pn - qn^2)$  for some constants  $p, q > 0$ . Then for any  $n > 1$  we have that

$$\sum_{k=n}^{\infty} |x_k| \leq \frac{\sqrt{\pi}}{2\sqrt{q}} \exp\left(\frac{p^2}{4q}\right) \exp\left(-q\left(n - \frac{p}{2q}\right)^2\right).$$

**Proof** The result follows by straightforward calculation using the classical bound for the error function  $\int_n^{\infty} \exp(-t^2) dt \leq \frac{\sqrt{\pi}}{2} \exp(-n^2)$ .  $\square$

<sup>7</sup> The most time-consuming part is the Newton method used to locate a zero starting from a point of a lattice on  $\mathcal{R}(T)$ . The time taken by this calculation grows exponentially with  $n$ . The total number of the searches is proportional to the area of  $\mathcal{R}(T)$ , which is proportional to  $T$ .



**Fig. 9** Plots of the zeros  $Z_n(s)$  for  $n = 2, 4, \dots, 14$  and  $b = 6$

We now turn to the proof of Theorem 4.2. This follows the same lines as [17]. However, the key new idea is that the disks  $U_1$ ,  $U_2$  and  $U_3$  used to define  $\mathcal{B}$  are allowed to depend on  $b$ .

**Proof of Theorem 4.2** Without loss of generality we may assume that the geodesic  $\beta_j$  has end points  $e^{(\frac{2\pi}{3}j \pm \theta)i} \in \partial \mathbb{D}^2$  for  $j = 1, 2, 3$ . We choose three additional geodesics  $v_j$  with end points  $e^{(\frac{2\pi}{3}j \pm \varphi)i} \in \partial \mathbb{D}^2$  for some  $0 < \varphi < \theta$ , that we will specify later, see Fig. 8 for details. We may consider  $\mathbb{D}^2$  as a subset of  $\mathbb{C}$  with usual Euclidean metric, and then complete  $v_j$  to full Euclidean circles  $\bar{v}_j \in \mathbb{C}$ . We define  $U_j \in \mathbb{C}$  to be compact disks with  $\partial U_j = \bar{v}_j$ .

It turns out that the calculations are much easier in the upper half plane model of the hyperbolic space  $\mathbb{H}^2$ . We choose the map  $S(z) = i \frac{1-z}{1+z}$  to change the coordinates. Then the geodesic  $\beta_j$  has end points  $\frac{\sin(\frac{2\pi}{3}j \pm \theta)}{1 + \cos(\frac{2\pi}{3}j \pm \theta)} \in \partial \mathbb{H}^2$ , and its Euclidean radius and centre are given by, respectively

$$\varepsilon_j = \frac{1}{2} \left( \frac{\sin(\frac{2\pi}{3}j + \theta)}{1 + \cos(\frac{2\pi}{3}j + \theta)} - \frac{\sin(\frac{2\pi}{3}j - \theta)}{1 + \cos(\frac{2\pi}{3}j - \theta)} \right) \quad (4.2)$$

$$c_j = \frac{1}{2} \left( \frac{\sin(\frac{2\pi}{3}j + \theta)}{1 + \cos(\frac{2\pi}{3}j + \theta)} + \frac{\sin(\frac{2\pi}{3}j - \theta)}{1 + \cos(\frac{2\pi}{3}j - \theta)} \right). \quad (4.3)$$

The end points of  $v_j$  are  $\frac{\sin(\frac{2\pi j}{3} \pm \varphi)}{1 + \cos(\frac{2\pi j}{3} \pm \varphi)} \in \partial \mathbb{H}^2$  and the euclidean radius is

$$r_j = \frac{1}{2} \left( \frac{\sin(\frac{2\pi j}{3} + \varphi)}{1 + \cos(\frac{2\pi j}{3} + \varphi)} - \frac{\sin(\frac{2\pi j}{3} - \varphi)}{1 + \cos(\frac{2\pi j}{3} - \varphi)} \right). \quad (4.4)$$

We can consider the Banach space  $\mathcal{B}$  to be the space of bounded analytic functions on  $\sqcup_{j=1}^3 U_j \subset \mathbb{C}$  with the supremum norm.

We see that the reflection  $R_j$  with respect to the geodesic  $\beta_j$  in  $\mathbb{H}^2$  is given by  $R_j(z) = \frac{\varepsilon_j^2}{z - c_j} + c_j$ . We deduce that for any distinct  $j, k, l$  the image  $R_j(U_k \cup U_l) \subset U_j$ , provided  $\frac{\varepsilon_j^2}{z - c_j} < r_j$  for any  $z \in U_k \sqcup U_l$ . We know that for all  $z \in U_k \sqcup U_l$  we have  $|z - c_j| > 1$ , thus it is sufficient to choose  $\theta$  and  $\varepsilon_j$  such that  $\varepsilon_j^2 < r_j$ .

Using (4.2) by a straightforward calculation we may estimate  $\frac{1}{2}\theta \leq \varepsilon_j \leq 2\theta + O(\theta^2)$  and similarly from (4.4) we have that  $\frac{1}{2}\varphi \leq r_j \leq 2\varphi + O(\varphi^2)$  for small values of  $\theta$  and  $\varphi$ . Hence it is sufficient to choose  $\theta$  and  $\varphi$  such that  $4\theta^2 < \frac{1}{2}\varphi$ . Using Lemma 3.4, we see

$$\theta = e^{-b}(1 + e^{-2b} + o(e^{-3b})).$$

In particular, it is sufficient to choose

$$\varphi = e^{-b\kappa} \text{ for some } 1 < \kappa < 2 \text{ and } b \text{ sufficiently large.} \quad (4.5)$$

Then (4.2) and (4.4) give estimates for the radii of inner and outer circles, respectively,

$$\frac{1}{2}e^{-b\kappa} \leq r_j \leq 2e^{-b\kappa} + O(e^{-2b\kappa}), \quad (4.6)$$

$$\frac{1}{2}e^{-b} \leq \varepsilon_j \leq 2e^{-b} + O(e^{-2b}). \quad (4.7)$$

Using the Cauchy integral formula for  $z \in U_k$  and  $R_j(z) \in U_j$  we can write

$$f(R_j(z)) = \frac{1}{2\pi i} \int_{\partial U_j} \frac{f(\xi)}{\xi - R_j(z)} d\xi$$

and thus

$$(R'_j(z))^s f(R_j(z)) = \frac{(R'_j(z))^s}{2\pi i} \int_{\partial U_j} \frac{f(\xi)}{\xi - R_j(z)} d\xi.$$

Since  $\mathcal{L}_s$  is a nuclear operator, the identity (3.1) should hold. More precisely, we may write

$$\begin{aligned} (\mathcal{L}_s f)(z) &= \sum_{j=1}^3 (R'_j(z))^s f(R_j(z)) \sum_{k=1, k \neq j}^3 \chi_{U_k}(z) \\ &= \sum_{j=1}^3 (R'_j(z))^s \cdot \left( \sum_{n=0}^{\infty} \frac{(R_j(z) - c_j)^n}{2\pi i} \int_{\partial U_j} \frac{f(\xi)}{(\xi - c_j)^{n+1}} d\xi \right) \cdot \left( \sum_{k=1, k \neq j}^3 \chi_{U_k}(z) \right) \\ &= \sum_{n=0}^{\infty} \lambda_n w_n(z) v_n(f), \end{aligned}$$

where,  $w_n \in \mathcal{B}$ ,  $v_n \in \mathcal{B}^*$ , and  $\lambda_n \in \mathbb{R}^+$  satisfy conditions of Definition 3.1. We may choose for any  $j \in \{1, 2, 3\}$

$$w_{3n+j}(z) \asymp (R'_j(z))^s \cdot \frac{(R_j(z) - c_j)^n}{2\pi i} \sum_{k=1, k \neq j}^3 \chi_{U_k}(z) \quad (4.8)$$

$$v_{3n+j}(f) \asymp \int_{\partial U_j} \frac{f(\xi)}{(\xi - c_j)^{n+1}} d\xi, \quad (4.9)$$

with normalization  $\|w_n\|_\infty = \|v_n\|_\infty = 1$ , and where  $c_j$  are given by (4.3). Then for any  $j \in \{1, 2, 3\}$

$$|\lambda_{3n+j}| = \left\| (R'_j)^s \right\|_{U_k \cup U_l} \cdot \left\| (R_j - c_j)^n \right\|_{U_k \cup U_l} \cdot \left\| \frac{1}{2\pi i} \int_{\partial U_j} \frac{f(\xi)}{(\xi - c_j)^{n+1}} d\xi \right\|_\infty, \quad (4.10)$$

where  $k \neq j$  and  $l \neq j$ , and the latter term is the supremum norm of the functional

$$f \mapsto \frac{1}{2\pi i} \int_{\partial U_j} \frac{f(\xi)}{(\xi - c_j)^{n+1}} d\xi.$$

We may observe that for any  $\xi \in \partial U_j$  one has that  $\xi - c_j = r_j$  and conclude

$$|\lambda_{3n+j}| \leq \|R'_j(z)^s\|_{U_k \cup U_l} \cdot \frac{\varepsilon_j^{2n}}{r_j^{n+1}}.$$

More precisely, using formulae (4.2) and (4.4), we obtain an upper bound

$$\begin{aligned} |\lambda_n| &\leq \max_j \|R'_j(z)^s\|_{U_k \cup U_l} \cdot \max \left\{ \frac{1}{r_1} \frac{\varepsilon_1^{2n/3}}{r_1^{n/3}}, \frac{1}{(\varepsilon_2 r_2)^{2/3}} \frac{\varepsilon_2^{2n/3}}{r_2^{n/3}}, \frac{1}{\varepsilon_3^{4/3} r_3^{1/3}} \frac{\varepsilon_3^{2n/3}}{r_3^{n/3}} \right\} \\ &\leq \max_j \|R'_j(z)^s\|_{U_k \cup U_l} \cdot \frac{1}{\varepsilon_3^{4/3} r_3^{1/3}} \max_j \frac{\varepsilon_j^{2n/3}}{r_j^{n/3}}. \end{aligned}$$

Comparing this with the definition of the nuclear operator 3.1, we get explicit bounds for parameters  $\lambda$  and  $C(s)$ .

$$|\lambda_n| \leq \max_k \sup_{z \in U_k} \max_{j \neq k} \left| \frac{\varepsilon_j^{2s}}{(z - c_j)^{2s}} \right| \cdot \frac{1}{\varepsilon_3^{4/3} r_3^{1/3}} \cdot \max_j \frac{\varepsilon_j^{2n/3}}{r_j^{n/3}} \leq C(s) \lambda^n,$$

with the choices

$$\lambda = \left( \max_j \frac{\varepsilon_j^2}{r_j} \right)^{\frac{1}{3}} \quad (4.11)$$

$$C(s) = \max_k \sup_{z \in U_k} \max_{j \neq k} \left| \frac{\varepsilon_j^{2s}}{(z - c_j)^{2s}} \right| \cdot \frac{1}{\varepsilon_3^{4/3} r_3^{1/3}}. \quad (4.12)$$

Using the bounds (4.6) and (4.7) for  $\varepsilon_j$  and  $r_j$ , we conclude

$$\lambda = \left( \max_j \frac{\varepsilon_j^2}{r_j} \right)^{\frac{1}{3}} \leq 2e^{-\frac{2-\varkappa}{3}b}. \quad (4.13)$$

Furthermore, we see that for any  $j \neq k$  for all  $z \in U_k$  we have  $|\arg(z - c_j)| \leq \arcsin(\frac{r_j}{c_k - c_j}) \leq 2\varphi$ . Therefore for  $s = \sigma + it$ ,

$$\begin{aligned} |(z - c_j)^{2s}| &= \left| \exp(2(\ln |z - c_j| + i \arg(z - c_j)) \cdot (\sigma + it)) \right| \\ &= |z - c_j|^{2\sigma} \cdot \exp(-2 \arg(z - c_j)t) \geq \exp(4\varphi t), \end{aligned}$$

since by construction  $\inf_{z \in U_k} |z - c_j| > 1$ . Using (4.6) and (4.7), we deduce

$$\begin{aligned} C(s) &= \max_k \sup_{z \in U_k} \max_{j \neq k} \left| \frac{\epsilon_j^{2s}}{(z - c_j)^{2s}} \right| \cdot \frac{1}{\epsilon_3^{4/3} r_3^{1/3}} \\ &\leq \max_k \sup_{z \in U_k} \max_{j \neq k} \left| \frac{\epsilon_j^{2\sigma}}{(z - c_j)^{2s}} \right| \cdot 4e^{b(\kappa+4)/3} \\ &\leq 4\epsilon_j^{2\sigma} e^{4\varphi t + b(\kappa+4)/3} \leq e^{\ln 4 - 2b\sigma + 4\varphi t + b(\kappa+4)/3}. \end{aligned} \quad (4.14)$$

Substituting bounds (4.13) and (4.14) into Ruelle's inequality (3.3) and taking into account  $t < T$  for  $s = \sigma + it \in \mathcal{R}(T)$ , we obtain an upper bound

$$\begin{aligned} |a_n(s)| &\leq C^n(s) \lambda^{n(n+1)/2} n^{n/2} \\ &\leq \exp \left( (\ln 4 - 2b\sigma + \frac{b(\kappa+4)}{3} + 4\varphi t)n - \frac{n(n+1)}{2} \left( \frac{b(2-\kappa)}{3} - \ln 2 \right) + \frac{n \ln n}{2} \right) \\ &\leq \exp \left( (\ln 4 + \frac{b(\kappa+4)}{3} + 4\varphi T)n - \frac{n(n+1)}{2} \left( \frac{b(2-\kappa)}{3} - \ln 2 \right) + \frac{n \ln n}{2} \right), \end{aligned} \quad (4.15)$$

since  $\exp(-2bn\sigma) \leq 1$ .

In order to estimate the tail of the series  $\sum_{n=14}^{\infty} a_n(s)$  using Lemma 4.5, it is sufficient to find a constant  $k_2 < 1$  such that

$$n \left( \ln 4 + \frac{b(\kappa+4)}{3} + \frac{b(2-\kappa)}{6} \right) + \frac{(n+1)n}{2} \ln 2 + \frac{n \ln n}{2} < \frac{bn^2(2-\kappa)}{6} k_2 \quad (4.16)$$

which is equivalent

$$\frac{\frac{15 \ln 2}{b} + 10 + \kappa}{k_2(2-\kappa) - \frac{3 \ln 2}{b}} < n - \frac{3 \ln n}{bk_2(2-\kappa) - 3 \ln 2}. \quad (4.17)$$

It is clear that the last inequality doesn't hold for any  $n \leq 10$  but it does hold, for example, for all  $b \geq 20$  and  $n \geq 14$  with the choices  $\kappa = 1.05$  and  $0.95 \leq k_2 < 1$ . Therefore we obtain an upper bound

$$|a_n(s)| \leq \exp \left( 4\varphi T n - \frac{b(2-\kappa)(1-k_2)}{6} n^2 \right), \quad \text{for all } n \geq 14. \quad (4.18)$$

We recall that  $\varphi = e^{-\kappa b}$  by (4.5) and applying Lemma 4.5 with the choices  $p = 4e^{-\kappa b} T$  and  $q = \frac{b(2-\kappa)(1-k_2)}{6}$ , we get an estimate  $\sum_{k=n}^{\infty} |a_n(\sigma + it)| \leq \eta(b, n, T(b))$ , where  $T(b) = k_0 e^{\kappa b}$  for some  $k_0 > 0$ , all  $b \geq 20$ ,  $n \geq 14$  and

$$\begin{aligned} \eta(b, n, T(b)) &= \frac{\sqrt{6\pi}}{2\sqrt{b(2-\varkappa)(1-k_2)}} \exp\left(\frac{24k_0^2}{b(2-\varkappa)(1-k_2)}\right) \\ &\times \exp\left(\frac{b(2-\varkappa)(k_2-1)}{6} \left(n - \frac{12k_0}{b(2-\varkappa)(1-k_2)}\right)^2\right). \end{aligned} \quad (4.19)$$

Therefore we have the desired asymptotic estimates:

1. for any  $n \geq 14$  we have  $\eta(b, n, T(b)) = O\left(\frac{1}{\sqrt{b}}\right)$  as  $b \rightarrow \infty$ ;
2. for any  $b \geq 20$  we have  $\eta(b, n, T(b)) = O\left(e^{-bk_1 n^2}\right)$  as  $n \rightarrow \infty$ ;

hold with the choices, for example,  $0 < k_1 \leq \frac{(2-\varkappa)(1-k_2)}{6}$ , where  $1 < \varkappa < 2$  and  $k_2$  are chosen so that (4.16) holds.  $\square$

## 5 Results on the zero set

We now turn to the problem of describing the distribution of the zeros. In Sect. 2.2 we introduced a matrix function  $A(s)$ , closely connected to the zeta function. In the following proposition we study the convergence of

$$\lim_{n \rightarrow \infty} \det(I_N - A^2(s));$$

from Lemma 2.8.

Let us recall the matrix  $A(s)$  computed in (2.16) using an approximation to the length of closed geodesics based on the segments of word length 2:

$$A^2(s) = e^{-2sb} \begin{pmatrix} 1 & z & 0 & 0 & z^2 & z \\ z & 1 & z^2 & z & 0 & 0 \\ 0 & 0 & 1 & z & z & z^2 \\ z^2 & z & z & 1 & 0 & 0 \\ 0 & 0 & z & z^2 & 1 & z \\ z & z^2 & 0 & 0 & z & 1 \end{pmatrix}, \quad \text{where } z = e^{-se^{-b}}.$$

As we are looking to study rescaled zeros,

$$\left\{ \sigma b + i e^{-b} t \mid Z_{X_b}(\sigma + it) = 0 \right\} = \left\{ \sigma + it \mid Z_{X_b}\left(\frac{\sigma}{b} + i t e^b\right) = 0 \right\},$$

it is appropriate to consider  $A\left(\frac{\sigma}{b} + i t e^b\right) = e^{-2\sigma - 2itbe^b} B(z)$ , where  $z = \exp\left(-\frac{1}{2}\left(\frac{\sigma}{be^b} + it\right)\right)$  and  $B(z) = e^{2bs} A^2(s)$ , defined by (2.17). Taking into account that  $\exp\left(-\frac{\sigma}{2be^b}\right) \rightarrow 1$  as  $b \rightarrow +\infty$  for  $\sigma > 0$ , we conclude the following:

**Proposition 5.1** *Using the notation introduced above, the real analytic function  $Z_{12}\left(\frac{\sigma}{b} + i t e^b\right)$  converges uniformly to  $\det(I - \exp(-2\sigma - 2itbe^b)B(e^{it}))$  on the critical strip, and more precisely,*

$$\left| Z_{12}\left(\frac{\sigma}{b} + i t e^b\right) - \det\left(I - \exp(-2\sigma - 2itbe^b)B(e^{it})\right) \right| = O\left(e^{-b}\right) \text{ as } b \rightarrow +\infty.$$

**Proof** This follows by straightforward calculation of the first 12 coefficients and the determinant. Let us introduce dummy variables  $x = \exp(-2\sigma - 2itbe^b)$  and  $y = e^{it}$  with  $|x| < 1$  and  $|y| = 1$ . Then

$$\det(I - \exp(-2\sigma - 2itbe^b)B(e^{it})) = \det(I - xB(y)) = \sum_{k=0}^6 x^k P_k(y^2), \quad (5.1)$$

where  $P_k \in \mathbb{Z}[\cdot]$  are some polynomials with integer coefficients. More precisely, we can compute:

$$\begin{aligned} P_0(y) &\equiv 1, \\ P_1(y) &\equiv 6, \\ P_2(y) &= 15 - 6y^2, \\ P_3(y) &= 20 - 24y^2 + 6y^4 + 2y^6, \\ P_4(y) &= 15 - 36y^2 + 27y^4 - 6y^6, \\ P_5(y) &= -6(y^2 - 1)^4, \\ P_6(y) &= (y^2 - 1)^6. \end{aligned}$$

On the other hand

$$Z_{12}\left(\frac{\sigma}{b} + ite^b\right) = \sum_{j=0}^{12} a_j \left(\frac{\sigma}{b} + ite^b\right).$$

Now one can deduce by comparing coefficients in  $x^n$  that  $a_{2n}\left(\frac{\sigma}{b} + ite^b\right) = x^n P_n(y) + O(e^{-b})$ .  $\square$

**Remark 5.2** Using the estimates for the hyperbolic length of closed geodesics of the word length  $\omega(\gamma) \leq 6$ , presented in the “Appendix A”, we may explicitly compute the first few non-zero coefficients

$$\begin{aligned} a_2(s) &= -\frac{6e^{-2bs}}{1 - e^{-2b}} \\ a_4(s) &= 15e^{-4bs} - 6e^{-4bs}(1 + e^{-b} + 2e^{-2b})^{2s} + O(e^{-b}) \\ a_6(s) &= -20e^{-6bs} - 2e^{-6bs}(1 + 3e^{-b} + 3e^{-2b})^{2s} - 6e^{-6bs}(1 + 2e^{-b} + 3e^{-2b})^{2s} + \\ &\quad + 24e^{-6bs}(1 + e^{-b} + 2e^{-2b})^{2s} + O(e^{-b}). \end{aligned}$$

We shall illustrate Proposition 5.1 using formulae for the coefficients above. We can write

$$a_2\left(\frac{\sigma}{b} + ite^b\right) = -\frac{6\exp(-2\sigma - 2itbe^b)}{1 - e^{-2b}} = \frac{-6x}{1 - e^{-2b}} = -6x(1 + O(e^{-2b})) = xP_1 + O(e^{-b}).$$

Similarly for  $a_4$ :

$$\begin{aligned} a_4\left(\frac{\sigma}{b} + ite^b\right) &= 15e^{-4\sigma - 4itbe^b} - 6e^{-4\sigma - 4itbe^b}(1 + e^{-b} + 2e^{-2b})^{2\sigma/b + 2ite^b} + O(e^{-b}) \\ &= 15x^2 - 6x^2(1 + e^{-b} + 2e^{-2b})^{2\sigma/b} \cdot (1 + e^{-b} + 2e^{-2b})^{2ite^b} + O(e^{-b}) \\ &= 15x^2 - 6x^2e^{2it} \cdot (1 + O(e^{-b})) + O(e^{-b}) \\ &= x^2(15 - 6y^2) + O(e^{-b}) = x^2P_2(y) + O(e^{-b}), \end{aligned}$$



where we have used the fact that  $(1 + e^{-b} + 2e^{-2b})^{e^b} = e + O(e^{-b})$ .

Now we can prove Theorem 1.5.

**Proof of Theorem 1.5** We find that the matrix  $B(e^{it})$  has exactly four different eigenvalues  $\mu_k(t)$ ,  $k = 1, \dots, 4$ :

$$\begin{aligned}\mu_1(t) &= (e^{it} - 1)^2 \\ \mu_2(t) &= (e^{it} + 1)^2 \\ \mu_3(t) &= 1 - \frac{e^{2it}}{2} + e^{it} \frac{\sqrt{4 - 3e^{2it}}}{2} \\ \mu_4(t) &= 1 - \frac{e^{2it}}{2} - e^{it} \frac{\sqrt{4 - 3e^{2it}}}{2}.\end{aligned}$$

By Lemma 2.11, the zero set of the determinant  $\det(I - \exp(-2\sigma - 2itbe^b)B(e^{it}))$  belongs to the subset  $\{(\sigma, t) \in \mathbb{R}^2 \mid \exists k: |\exp(2\sigma + 2itbe^b)| = |\mu_k(t)|\}$ . The four equations  $\exp(2\sigma) = |\mu_k(t)|$  give us four curves

$$\begin{aligned}\mathcal{C}_1 &= \left\{ \frac{1}{2} \ln |2 - 2\cos(t)| + it \mid t \in \mathbb{R} \right\}; \\ \mathcal{C}_2 &= \left\{ \frac{1}{2} \ln |2 + 2\cos(t)| + it \mid t \in \mathbb{R} \right\}; \\ \mathcal{C}_3 &= \left\{ \frac{1}{2} \ln \left| 1 - \frac{1}{2}e^{2it} - \frac{1}{2}e^{it}\sqrt{4 - 3e^{2it}} \right| + it \mid t \in \mathbb{R} \right\}; \\ \mathcal{C}_4 &= \left\{ \frac{1}{2} \ln \left| 1 - \frac{1}{2}e^{2it} + \frac{1}{2}e^{it}\sqrt{4 - 3e^{2it}} \right| + it \mid t \in \mathbb{R} \right\}.\end{aligned}$$

Since the curves  $\mathcal{C}_j$  do not have horizontal tangencies  $\sigma \equiv \text{const}$ , without loss of generality we may define neighbourhoods as follows:

$$V(\mathcal{C}_j, \varepsilon) = \{(\sigma, t) \mid |2\sigma - \ln |\mu_j(t)|| < 2\varepsilon\}.$$

To complete the argument we shall show that for all  $\varepsilon > 0$  and  $T > 0$  there exists  $b_0 > 0$  such that for any  $b > b_0$  the zeros of the function  $Z\left(\frac{\sigma}{b} + ite^b\right)$  with  $0 \leq \sigma \leq 1$  and  $|t| \leq e^{(2-\varepsilon)b}$  belong to a neighbourhood  $\cup_k V(\mathcal{C}_k, \varepsilon)$  of the union of the curves  $\cup_k \mathcal{C}_k$ .

Indeed, given  $\varepsilon > 0$  and a point  $z_0 = \sigma_0 + it_0$  outside of  $\varepsilon$ -neighbourhood of  $\cup_{j=1}^4 \mathcal{C}_j$  we see that the determinant

$$\left| \det \left( I - \exp \left( -2\sigma_0 - it_0be^b \right) B(\exp(it_0)) \right) \right| > \exp(-6\varepsilon)(\exp \varepsilon - 1)^6 > 0$$

is bounded away from zero and the bound is independent of  $b$ . Summing up, we see that outside of the neighbourhood  $\cup_{j=1}^4 V(\mathcal{C}_j, \varepsilon)$  the determinant has modulus uniformly bounded away from 0; by Theorem 4.2 for  $b$  large we have that the zeta function  $Z_{X_b}\left(\frac{\sigma}{b} + ite^b\right)$  can be approximated by  $Z_{12}$  arbitrarily closely and by Proposition 5.1  $Z_{12}$  can be approximated arbitrarily closely by the determinant. Therefore for  $b$  sufficiently large all zeros of the function  $Z_{X_b}\left(\frac{\sigma}{b} + ite^b\right)$  belong to the  $\varepsilon$ -neighbourhood of  $\cup_{j=1}^4 \mathcal{C}_j$ .  $\square$

We have concentrated on the particular case of the symmetric 3-funnelled surface (whose defining closed geodesics have the same lengths). However, the same method of combining geometric and analytic approximations works in the case that the boundary curves have different length as well as in the case of symmetric punctured torus, and allows one to explain the nature of the patterns of zeros described in the sections 5.1 and 5.2 of [6].

## 6 L-functions and covering surfaces

Our results have concentrated on a special class of surfaces, but can be easily adapted to cover a large class of geometrically finite surfaces of infinite area.

There is a fairly simple method for constructing quite complicated surfaces using any (infinite area) surface  $V$  without cusps. We can write  $V = \mathbb{D}^2 / \Gamma$  for a convex cocompact group  $\Gamma$ . Then we can define a (finite) cover  $\widehat{V} = \mathbb{H}^2 / \Gamma_0$  for  $V$  in terms of a (finite index) normal subgroup  $\Gamma_0 < \Gamma$ .

Let us denote by  $G = \Gamma / \Gamma_0$  the finite quotient group. Let  $\gamma$  be a closed geodesic on  $X_b$  and then this is covered by the union of closed geodesics  $\gamma_1, \dots, \gamma_n$  on  $\widehat{V}$ .

Let  $R_\chi$  be an irreducible representation for  $G$  of degree  $d_\chi$  with character  $\chi = \text{tr}(R_\chi)$ . The regular representation of  $G$  can be written  $R = \bigoplus_\chi d_\chi R_\chi$ , where  $|G| = \sum_\chi d_\chi^2$ .

**Definition 6.1** Given  $s \in \mathbb{C}$  we define

$$L(z, s, \chi) = \prod_\gamma \det \left( I - z^{|g|} e^{-(s+n)\lambda(g)} R(g\Gamma_0) \right)$$

where  $g\Gamma_0$  is a coset in  $G$ .

**Lemma 6.2** For characters  $\chi_1$  and  $\chi_2$  we can write

$$L(z, s, \chi_1 + \chi_2) = L(z, s, \chi_1) L(z, s, \chi_2).$$

If  $H < G$  is a subgroup and  $\chi$  is a character of  $H$  then we can write  $G = \bigcup_{i=1}^m H\alpha_i$  and define the induced character  $\chi^*$  of  $G$  by

$$\chi^*(g) = \sum_{\alpha_i g \alpha_i^{-1} \in H} \chi(\alpha_i g \alpha_i^{-1})$$

for  $g \in G$ .

**Lemma 6.3** (Brauer–Frobenius) Each non-trivial character  $\chi$  is a rational combination of characters  $\chi_i^*$  of  $G$  induced from non-trivial characters  $\chi_i$  of cyclic subgroups  $H_i$ .

There exist integers  $n_1, \dots, n_k$  with

$$n\chi = \sum_{i=1}^k n_i \chi_i^*$$

and thus

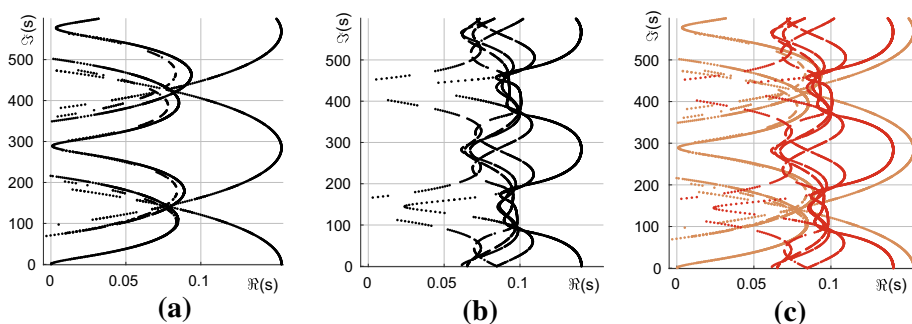
$$L(s, z, \chi)^n = \sum_{\alpha_i g \alpha_i^{-1} \in H} \chi(\alpha_i g \alpha_i^{-1}).$$

Since it is easier to deal with cyclic covering groups. We need the following.

**Lemma 6.4** Let  $\chi$  be a character of the subgroup  $H < G$  and let  $\widehat{L}(s, z, \chi^*)$  be the L-function with respect to the covering  $\widehat{V}$  of  $\widehat{V}/H$ . Then  $L(s, z, \chi) = \widehat{L}(s, z, \chi^*)$ .

The proof is analogous to that of the proof of Proposition 2 in [24].

This leads to the following.



**Fig. 10** (a) Zeros of the zeta function (b) zeros of the  $L$ -function in the case  $\ell(\gamma_{1,2,3} = 9)$ . (c) Superposition of (a) and (b). The apparent gaps are due to instability of the Newton method

**Lemma 6.5** *If  $\chi$  is an irreducible non-trivial character of  $G$  then  $L(s, \chi)^n$  is a product of integer powers of  $L$ -functions defined with respect to non-trivial characters of cyclic subgroups of  $G$ .*

Finally this means that we can write the zeta function  $Z_{\widehat{V}}(s, z)$  in terms of the  $L$ -functions  $L(s, z, \chi)$  for  $V$ .

**Lemma 6.6** *We can write*

$$Z_{\widehat{V}}(s, z) = \prod_{\chi \text{ irreducible}} L(s, z, \chi)^{d_{\chi}}.$$

where the product is over all irreducible representations of  $G$ .

In particular, the zeros for  $Z_{\widehat{V}}(s, z)$  will be the union of the zeros for the  $L$ -functions  $L_V(s, z, \chi)^{d_{\chi}}$ .

**Example 6.7** We can take a double cover  $\widehat{X}_b$  for a three funnelled surface  $X_b$ , which corresponds to a 4-funnelled surface. The corresponding covering group is simply  $\mathbb{Z}_2$  and the zeta function  $Z_{\widehat{X}_b}(s)$  is then the product of:

1. the zeta function  $Z_{X_b}(s)$  for the original surface;
2. the  $L$ -function  $L(s, \chi) := L(s, 1, \chi)$  corresponding to the representation  $\chi : \pi_1(X_b) \rightarrow \mathbb{Z}_2$  where  $\chi(g) = (-1)^{n(g)}$  where  $n(g)$  counts the number of times the generator  $a$ , say, occurs in  $g$ .

In particular the zeros for  $Z_{\widehat{X}_b}(s)$  are a union of the figures for these two functions (see Fig. 10 for a plot).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix A: Examples of the coefficients

In this Appendix we present the asymptotic formulae for the hyperbolic length of the short closed geodesics, which then lead to the asymptotic expressions for the first few non-zero coefficients  $a_2, a_4, a_6$ .

Using the identity

$$\ell(\gamma_{j_1, j_2, \dots, j_{2n}}) = 2\text{Arcosh}\left(\frac{1}{2}\text{tr}(R_{j_1} R_{j_2} \dots R_{j_{2n}})\right),$$

relating the length of the closed geodesic corresponding to the cutting sequence of period  $2n$  to the matrices defining the reflections, we compute the lengths of the closed geodesics for  $n = 2, 4, 6, 8$ .

**The case**  $n = 2$ . It has been established in Lemma 2.10, see also Remark 2.6, that there are exactly 6 geodesics of length  $\ell(\gamma_{j_1 j_2}) = 2b$ .

**The case**  $n = 4$ . There are 6 geodesics of length  $4b$  and 12 geodesics of length

$$\ell(\gamma_{j_1 j_2 j_3 j_2}) = 2\text{Arcosh}(\cosh(b) + 2\cosh^2(b)).$$

**The case**  $n = 6$ . There are  $4^3 + 2 = 66$  homotopy classes of closed geodesics; among which there are 6 geodesics of length  $6b$  and of length

$$\ell(\gamma_{j_1 j_2 j_3 j_1 j_2 j_3}) = 2\text{Arcosh}(4\cosh^3(b) + 6\cosh^2(b) - 1) = 6b + 6e^{-b} + O(e^{-2b}).$$

There are 18 geodesics of length

$$\ell(\gamma_{j_1 j_2 j_1 j_3 j_2 j_3}) = 2\text{Arcosh}\left(8\cosh^2\left(\frac{b}{2}\right) \cdot \cosh^2(b) - 1\right) = 6b + 4e^{-b} + O(e^{-2b}).$$

Finally, there are 36 geodesics of length

$$\ell(\gamma_{j_3 j_2 j_1 j_2 j_1 j_2}) = 2\text{Arcosh}(4\cosh^3(b) + 2\cosh^2(b) - \cosh(b)) = 6b + 2e^{-b} + O(e^{-2b}).$$

**The case**  $n = 8$ . There are  $4^4 + 2 = 258$  homotopy classes of closed geodesics; among which there are 6 geodesics of length  $8b$ . Moreover, there are 24 geodesics of length

$$\begin{aligned} \ell(\gamma_{j_3 j_1 j_3 j_1 j_2 j_1 j_2 j_1}) &= 2\text{Arcosh}(-4\cosh^2(b) + 4\cosh^3(b) + 8\cosh^4(b) + 1) \\ &= 8b + 2e^{-b} + O(e^{-2b}); \end{aligned}$$

and another 48 geodesics of length

$$\begin{aligned} \ell(\gamma_{j_3 j_1 j_2 j_1 j_2 j_1 j_2 j_1}) &= 2\text{Arcosh}(-\cosh(b) - 4\cosh^2(b) + 4\cosh^3(b) + 8\cosh^4(b)) \\ &= 8b + 2e^{-b} + O(e^{-2b}). \end{aligned}$$

In addition, we have 12 geodesics of length

$$\begin{aligned} \ell(\gamma_{j_3 j_1 j_2 j_1 j_3 j_1 j_2 j_1}) &= 2\text{Arcosh}(2\cosh^2(b) + 8\cosh^3(b) + 8\cosh^4(b) - 1) \\ &= 8b + 4e^{-b} + O(e^{-2b}); \end{aligned}$$

and 48 geodesics of length

$$\begin{aligned} \ell(\gamma_{j_3 j_1 j_2 j_3 j_1 j_2 j_1 j_2}) &= 2\text{Arcosh}(\cosh(b) \cdot (4\cosh(2b) + 2\cosh(3b) + 4\cosh(b) + 1)) \\ &= 8b + 4e^{-b} + O(e^{-2b}); \end{aligned}$$

and another 48 geodesics of length

$$\begin{aligned}\ell(\gamma_{j_3 j_2 j_1 j_3 j_2 j_1 j_2}) &= 2\text{Arcosh}(-\cosh(b) + 8\cosh^3(b) + 8\cosh^4(b)) \\ &= 8b + 4e^{-b} + O(e^{-2b}).\end{aligned}$$

Finally, there are 24 geodesics of length

$$\begin{aligned}\ell(\gamma_{j_3 j_2 j_1 j_2 j_3 j_1 j_2 j_1}) &= 2\text{Arcosh}(\cosh(b) \cdot (6\cosh(2b) + 2\cosh(3b) + 8\cosh(b) + 3)) \\ &= 8b + 6e^{-b} + O(e^{-2b});\end{aligned}$$

and further more 48 geodesics of length

$$\begin{aligned}\ell(\gamma_{j_3 j_1 j_2 j_3 j_1 j_2 j_1 j_2}) &= 2\text{Arcosh}(-3\cosh(b) + 12\cosh^3(b) + 8\cosh^4(b)) \\ &= 8b + 6e^{-b} + O(e^{-2b}).\end{aligned}$$

## Appendix B: Non-periodicity

The apparent almost periodicity in the plot can never be exact for a fixed  $b$  as we see from the behaviour of the zeros near the vertical line  $\Re(s) = \delta$  in Fig. 11. In fact, since the geodesic flow restricted to the non-wandering set is mixing, it is shown in [25] that there is only one zero with  $\Re(s) = \delta$ . Naud [22] (see also Jacobson & Naud [16]) showed an even stronger result: there exists  $\varepsilon > 0$  such that there is only finite number of zeros satisfying  $\Re(s) > \delta - \varepsilon$ . This is illustrated by the numerical results in Table 1. Namely, we analyze values of zeros closest to the right boundary  $\Re(z) = \delta$  of the critical strip:

$$\mathcal{E} := \left\{ s_0 \in \mathcal{S}_{X_b} \mid \text{for all } s \in \mathcal{S}_{X_b} \text{ such that } |s - s_0| < \frac{\pi}{2}e^b \text{ we have } \Re(s) < \Re(s_0) \right\}.$$

We see that for all  $s \in \mathcal{E}$  satisfying  $\Im(s) < 10^3$  there exist an  $s' \in \mathcal{E}$  such that<sup>8</sup>

$$|s - s' + \pi e^b| \leq 3.$$

Apparently, related results have been observed in [7].

## Appendix C: Spacing of zeros

For completeness, in this Appendix we describe a slight strengthening of a particular case of a result of Weich on the spacing of imaginary parts of zeros. The Theorem below asserts that the spacing of the zeros  $\mathcal{S}_{X_b}$  for the zeta function is approximately  $\frac{\pi}{b}$  as  $b \rightarrow +\infty$ .

**Notation C.1** A compact part of the critical strip of width  $\delta = \delta(b) > 0$  and height  $T$  which we denote by

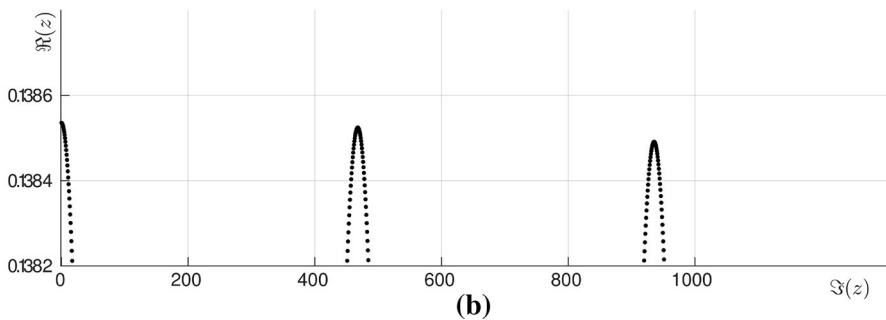
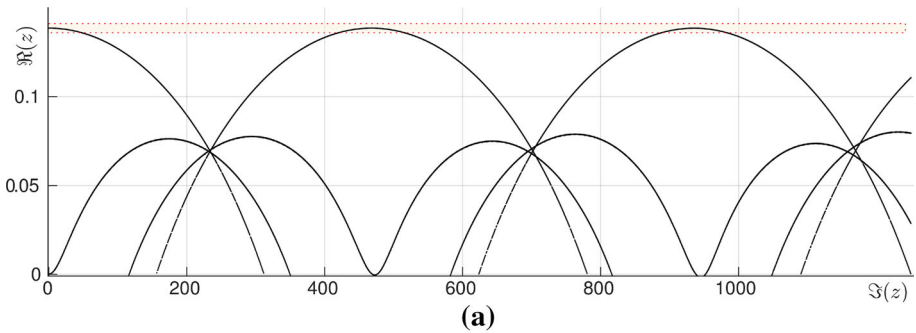
$$\mathcal{R}_b(T) = \{s \in \mathbb{C} \mid 0 \leq \Re(s) \leq \delta \text{ and } |\Im(s)| \leq T\}.$$

We denote a set of regularly spaced points on the lines  $\Re(s) = 0$  and  $\Re(s) = \ln 2$  given by  $\mathcal{L} = (\ln 2 + i\pi)\mathbb{Z} \cup i\pi\mathbb{Z}$ .

<sup>8</sup> We have, in fact, verified this for larger values of  $\Im(s)$ , but we omit the numerics here.

**Table 1** Empirical estimates on zeros near the right boundary  $\Re(z) = \delta$ 

Endpoints for strings of zeros with $\Re(z) \approx \delta$		
$k$	$z_k$	$z_k - z_{k-1}$
$2b = 3\pi, z_0 = \delta = 0.146949, \pi \exp(b) = 349.715115$		
1	$0.146928 + i351.330281$	$2.093 \times 10^{-5} + i351.33028$
2	$0.146866 + i702.660561$	$6.278 \times 10^{-5} + i351.33028$
3	$0.146761 + i1053.990842$	$1.047 \times 10^{-4} + i351.33028$
$2b = 8, z_0 = \delta = 0.172887, \pi \exp(b) = 171.525147$		
1	$0.172785 + i172.781$	$-1.0196 \times 10^{-4} + i172.781053$
2	$0.172481 + i346.345$	$-3.0451 \times 10^{-4} + i173.564643$
3	$0.171974 + i519.126$	$-5.0674 \times 10^{-4} + i172.781053$
4	$0.171262 + i691.907$	$-7.1224 \times 10^{-4} + i172.781053$
5	$0.170343 + i865.472$	$-9.1839 \times 10^{-4} + i173.564643$
6	$0.169219 + i1038.253$	$-11.2437 \times 10^{-4} + i172.781054$


**Fig. 11** **a** A plot of the zeros in the case  $\ell(\gamma_0) = 10$ ; and **b** A scaled up version to see the apparent periodicity near  $\Re(s) = \delta$

**Theorem C.2** *The sets  $\mathcal{S}_{X_b}$  and  $\mathcal{L}$  are close in the Hausdorff metric on a small part of the critical strip. More precisely, there exists  $\varkappa > 1$  such that*

$$\text{dist}_H(b \cdot (\mathcal{S}_{X_b} \cap \mathcal{R}_b(\varkappa)), \mathcal{L}) = O\left(\frac{1}{\sqrt{b}}\right), \text{ as } b \rightarrow +\infty.$$

An earlier version of Theorem C.2 was established by Weich [33], where he also considered funnels whose widths are  $\mathbb{Z}$ -multiples of  $b$ . On the other hand, his results apply only on a bounded domain and without the error term.

In order to prove Theorem C.2 we need the following approximation Lemma.

**Lemma C.3** *Given a  $\varkappa > 1$  as in Theorem 4.2, the complex analytic function  $Z_{X_b}(\frac{s}{b})$  converges uniformly to  $Z_6(\frac{s}{b})$  on the domain  $\mathcal{R}(\varkappa b)$ , more precisely,*

$$\sup_{s \in \mathcal{R}(\varkappa b)} \left| Z_{X_b}\left(\frac{s}{b}\right) - Z_6\left(\frac{s}{b}\right) \right| = O\left(\frac{1}{\sqrt{b}}\right) \text{ as } b \rightarrow +\infty.$$

**Proof** By a straightforward manipulation using the lengths of the closed geodesics estimated in Appendix A we show that

$$a_n(s) = \exp(-nbs - 2b)(1 + O(\exp(-2b)))^s \text{ for } n = 8, 10, 12$$

and therefore

$$a_n\left(\frac{s}{b}\right) = \exp(-ns - 2b)(1 + O(\exp(-2b)))^{s/b} \rightarrow 0 \text{ as } b \rightarrow \infty.$$

The result follows from Theorem 4.2 where the corresponding terms for  $n = 8, 10$ , and  $12$  are of order  $\left(\frac{1}{\sqrt{b}}\right)$ .  $\square$

Now we are ready to prove Theorem C.2, which is easier than Theorem 1.5, because we can use complex analysis.

**Proof** We shall show that on the domain  $\mathcal{R}(\varkappa b)$  we have that the function  $Z_{X_b}(s)$  vanishes at  $s_n(b)$  such that  $\lim_{b \rightarrow \infty} s_n(b) \cdot b = (\ln 2 + i\pi n)$ .

The function  $\det(I - e^{-2s} A^2)$  vanishes at  $\{i\pi n, \ln 2 + i\pi n\}$ , for  $n \in \mathbb{Z}$ . For any sufficiently small  $\eta > 0$  we have that the closed balls

$$\overline{U(\ln 2 + 2i\pi n, \eta)} = \{s \in \mathbb{C} : |s - (\ln 2 + 2i\pi n)| \leq \eta\}$$

contains no more zeros. Let us denote  $\varepsilon = \inf_{s \in \partial U} |\det(I - e^{-2s} A^2)| > 0$ . Using Theorem C.3, we can now choose  $b$  sufficiently large so that we have

$$\inf_{s \in \partial U} \left| Z_{X_b}\left(\frac{s}{b}\right) - \det(I - e^{-2s} A^2) \right| < \frac{\varepsilon}{2}.$$

It then follows by Rouché's Theorem [1] that for any  $n \in \mathbb{N}$  the function  $Z_{X_b}(s)$  has exactly one zero  $s_n(b)$ , satisfying  $|s_n(b) - \frac{1}{b}(\ln 2 + i\pi n)| < \eta$ .  $\square$

This implies the asymptotic spacing of imaginary parts of zeros.

## References

1. Ahlfors, L.V.: Complex analysis. An Introduction to the Theory of Analytic Functions of One Complex Variable. International Series in Pure and Applied Mathematics, 3rd edn. McGraw-Hill Book Co., New York (1978)

2. Bandtlow, O.F., Jenkinson, O.: On the Ruelle eigenvalue sequence. *Ergod. Theory Dyn. Syst.* **28**(6), 1701–1711 (2008)
3. Beardon, A.F.: *The Geometry of Discrete Groups*. Corrected Reprint of the 1983 Original. Graduate Texts in Mathematics, vol. 91. Springer, New York (1995)
4. Berger, M.: *Lectures on Geodesics Riemannian Geometry*, Lectures on mathematics and physics. Mathematics, 33, (Tata Institute of Fundamental Research, Mumbai (1965)
5. Bohr, H.: *Almost Periodic Functions*. Chelsea Publishing Company, New York (1947)
6. Borthwick, D.: Distribution of resonances for hyperbolic surfaces. *Exp. Math.* **23**(1), 25–45 (2014)
7. Borthwick, D., Weich, T.: Symmetry reduction of holomorphic iterated function schemes and factorization of Selberg zeta functions. *J. Spectr. Theory* **6**, 267–329 (2016)
8. do Carmo, M.: *Riemann Geometry*. Birkhauser, Basel (1992)
9. Cvitanović, P., Eckhardt, B.: Periodic-orbit quantization of chaotic systems. *Phys. Rev. Lett.* **63**(8), 823–826 (1989)
10. Fathi, A., Laudenbach, F., Poenaru, V.: Travaux de Thurston sur les surfaces, *Astérisque*, 66–67. Soc. Math. France, Paris (1979)
11. Grothendieck, A.: Produits tensoriels topologiques et espaces nucléaires (French). *Mem. Amer. Math. Soc.* **16**, 140 (1955)
12. Grothendieck, A.: La théorie de Fredholm. *Bull. Soc. Math. France* **84**, 319–384 (1956)
13. Guillopé, L., Lin, K.K., Zworski, M.: The Selberg zeta function for convex co-compact Schottky groups. *Commun. Math. Phys.* **245**(1), 149–176 (2004)
14. Hejhal, D.: *The Selberg Trace Formula for  $PSL(2, \mathbb{R})$* , Lecture Notes in Mathematics 548. Springer, Berlin (1976)
15. Jakobson, D., Naud, F.: On the critical line of convex co-compact hyperbolic surfaces. *Geom. Funct. Anal.* **22**(2), 352–368 (2012)
16. Jakobson, D., Naud, F.: Resonances and density bounds for convex co-compact congruence subgroups of  $SL_2(\mathbb{Z})$ . *Isr. J. Math.* **213**(1), 443–473 (2016)
17. Jenkinson, O., Pollicott, M.: Calculating Hausdorff dimensions of Julia sets and Kleinian limit sets. *Am. J. Math.* **124**(3), 495–545 (2002)
18. Jessen, B.: Some aspects of the theory of almost periodic functions. In: Noordhoff, E.P., Groningen, N.V. (eds.) 1957 Proceedings of the International Congress of Mathematicians, Amsterdam, vol. 1, pp. 305–314. North-Holland Publishing Co., Amsterdam (1954)
19. Levin, B. J.: *Distribution of Zeros of Entire Functions*. Revised edition. Translations of Mathematical Monographs, 5. American Mathematical Society, Providence, RI (1980)
20. Mayer, D.: The thermodynamic formalism approach to Selberg’s zeta function for  $PSL(2, \mathbb{Z})$ . *Bull. Am. Math. Soc. (N.S.)* **25**(1), 55–60 (1991)
21. McMullen, C.T.: Hausdorff dimension and conformal dynamics. III. Computation of dimension. *Am. J. Math.* **120**(4), 691–721 (1998)
22. Naud, F.: Expanding maps on Cantor sets and analytic continuation of zeta functions. *Ann. Sci. Ecole Norm. Sup.* **38**, 116–153 (2005)
23. Parry, W.: An analogue of the prime number theorem for closed orbits of shifts of finite type and their suspensions. *Isr. J. Math.* **45**, 41–52 (1983)
24. Parry, W., Pollicott, M.: The Chebotarev theorem for Galois coverings of Axiom A flows. *Ergod. Theory Dyn. Syst.* **6**, 133–148 (1986)
25. Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*. **187–188**, 1–268 (1990)
26. Patterson, S.J., Perry, P.A.: The divisor of Selberg’s zeta function for Kleinian groups. *Duke Math. J.* **106**(2), 321–390 (2001)
27. Pollicott, M.: Some applications of thermodynamic formalism to manifolds with constant negative curvature. *Adv. Math.* **85**(2), 161–192 (1991)
28. Ruelle, D.: Zeta-functions for expanding maps and Anosov flows. *Invent. Math.* **34**(3), 231–242 (1976)
29. Selberg, A.: Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Indian Math. Soc. (N.S.)* **20**, 47–87 (1956)
30. Series, C.: Geometrical Markov coding of geodesics on surfaces of constant negative curvature. *Ergod. Theory Dyn. Syst.* **6**, 601–625 (1986)
31. Sridhar, S., Lu, W.T.: Sinai billiards, Ruelle zeta-functions and Ruelle resonances: microwave experiments. *J. Stat. Phys.* **108**(5–6), 755–765 (2002)
32. Terras, A.: *Zeta Functions of Graphs. A Stroll Through the Garden*. Cambridge Studies in Advanced Mathematics, vol. 128. Cambridge University Press, Cambridge (2011)
33. Weich, T.: Resonance chains and geometric limits on Schottky surfaces. *Commun. Math. Phys.* **337**(2), 727–765 (2015)